PHYSIOLOGICAL REVIEWS

VOL. 1 JANUARY, 1921 No. 1

THE ORIGIN AND CONDUCTION OF THE HEART BEAT

J. A. E. EYSTER AND WALTER J. MEEK

From the University of Wisconsin, Madison

In the modern conception of the physiology of the heart beat, two contributions have played perhaps the most important part. The first of these is the proof that the activity of the heart may be analyzed into a number of different physiological properties or functions. The more important of these are automaticity, conductivity and contractility; properties too well recognized today to require definition or explanation. These properties are developed in different degrees in different parts of the heart. The heart beat arises in a relatively small area which, because it possesses the property of automaticity to the highest degree, initiates impulses for the rest of the heart. These impulses pass to all parts of the heart by virtue of the property of conductivity and cause contraction of the muscle fibers of the chambers. Removal of the influence of the region of highest automaticity by its destruction, by its functional separation, or finally by reduction of its automatic power by cooling or other means, results in some other region of lower automaticity assuming the function of impulse initiation.[1] Under these circumstances the original "pacemaker" is quiescent or continues to beat at its original rate but without influence upon the remainder of the heart. This fundamental conception was developed by work on the cold-blooded heart through the pioneer experiments of Stannius and the brilliant extension and interpretation of these by Gaskell and Englemann. According to Gaskell and Englemann, all parts of the heart have automaticity but in different degrees, this property being most developed

[1] Why the most automatic region should dominate the remainder of the organ and keep in abeyance the automatic power of other regions will probably not be entirely clear until the nature of the "inner stimulus" and the processes that underlie the initiation of the impulse are understood. The fact is that automatic power in any region only becomes manifest when removed from the influence of impulses coming to this region at a more rapid rate than the region itself is capable of initiating.

at the venous end, and least within the ventricles. The seat of impulse initiation normally resides within the sinus venosus. Gaskell, correllating this difference in rhythmic power in different regions with the embryonic development, regards it as most pronounced in those regions which retain to the greatest degree the "embryonic characteristics" and are least specialized for the purposes of contraction and the propulsion of blood.

The other contribution which we regard as fundamental for the present conception of the heart beat, is the series of physiological observations which indicates that recently discovered anatomical structures in the heart have the function of initiating and distributing the impulse on excitation to the contractile muscle tissues of the auricles and ventricles. Since much of the work contributary to this conception of cardiac activity is of comparatively recent date, we consider it proper to discuss it in some detail under separate headings.

ANATOMICAL CONSIDERATIONS

The conception of the origin and conduction of the heart beat in mammals now most widely prevalent, is intimately associated with certain recent histological observations, which tend to support the view of an anatomical as well as a physiological separation of those different properties of cardiac tissue, namely, automaticity, conductivity and contractility. The first of these observations was the discovery of His (1) that a separate bundle of tissue crosses the auriculo-ventricular junction in the mammalian heart. The proof of its functional importance for conduction of the impulse was carried out by His (2), Humblet (3), Hering (4) and especially Erlanger (5). A new realization of the importance of the type of tissue comprising this structure developed, however, from the observations of Tawara and of Keith and Flack. The histological observations of Tawara (6) showed that the bundle is composed of tissue which differs from the usual type of cardiac muscle, and that it forms a part of an extensive system connecting the auricles with the ventricles. The bundle begins above in a network of interlacing fibers of peculiar structure, lying near the base of the interauricular septum above the middle cusp of the tricuspid valve (the auriculo-ventricular node), and ends below in an extensive branching system reaching all parts of the endocardial surface of the two ventricles. The terminal branches are composed of large cells with prominent nuclei, the previously known Purkinje cells. The auriculo-ventricular node shows marked histological differentiation from the usual muscu-

lature of the auricles and ventricles, and its discovery presented an entirely new contribution to our knowledge of the more intimate structure of the heart.

The presence of a similar collection of tissue in the sulcus terminalis between the superior vena cava and the right auricle (the sino-auricular node), was discovered by Keith and Flack (7) in 1906. This tissue had been apparently observed previously by Wenckebach (8), who however interpreted it as a fibrous separating wall with a small bundle crossing it connecting the superior vena cava with the right auricle. Its structural identity with the auriculo-ventricular node was appreciated by Keith and Flack, and these workers regarded it as forming an important remains of the sinus venosus, and therefore probably important from the functional standpoint of impulse initiation. The main histological elements of both the auriculo-ventricular and sino-auricular nodes are slender, interlacing fibers, which stain lightly, contain many elongated nuclei and are imbedded in closely packed connective tissue. The sarcoplasm is relatively large in amount and the fibrillae much less numerous than in the usual cardiac muscle fiber. Nuclei are present in large numbers and nerve fibers and ganglia are abundant within and near the node (9), (10). Both nodes also receive a special and profuse blood supply.[2]

The anatomical characteristics of the nodes, the distribution of the ventricular system of fibers and the exact position and extent of the two nodes have been studied by numerous subsequent workers. The observations quoted above have been abundantly confirmed and many other important facts added. Observations from the developmental standpoint, by embryological studies made on a single species or by comparative studies in different species, have especial interest in relation to the physiological importance and interpretation of the function of these structures. It is beyond the scope of this paper to consider this work in detail, but certain conclusions that may be drawn from the work of Koch, (12), (13), (14), (15), (16), Keith (17), Keith and Mackenzie (18), Mackenzie (19), (20), Külbs (21), Aschoff (22) and others, based in part on the older embryological observations of His and Born, would seem to be of sufficient importance to warrant presentation.

[2] In further support of the anatomical and functional similarity of the two nodes, there may be mentioned the presence of an unusually high glycogen content in each, (11) and the fact, noted by Koch (12), that neither node takes part in hypertrophy or atrophy affecting the usual cardiac muscle of the auricles and ventricles.

1. The sino-auricular node and a part at least of the auriculo-ventricular node are derived from the primitive sinus venosus. The auriculo-ventricular node is divided into two parts, an auricular portion, or "atrial node" and a ventricular portion (22), (16). The former is probably derived, along with the sino-auricular node, from the primitive sinus valves, while the latter arises from the auricular canal (23). The atrial node surrounds the opening of the coronary sinus and is sometimes designated as the coronary sinus portion of the auriculo-ventricular node.

2. The auriculo-ventricular bundle and probably a part of the auriculo-ventricular node are derived from the auricular canal of the primitive heart. In the lower adult reptilian heart, and in the early mammalian embryo, the auricular canal forms a prominent connection running from the junction of the sinus and atrium to the ventricle. The auricle is in large part a side growth from this canal. There is present a muscular path connecting sinus and ventricle which, according to Mall (24), can be still recognized in the 8 mm. embryo as a broad strand connecting these two regions.

3. In the adult hearts of amphibia and reptiles, the auriculo-ventricular connection consists of two broad circular bands of vertically coursing fibers, forming really a right and left bundle, passing down from the lower border of the auricles deeply into the ventricle, where connections are made with the ventricular muscle. The separation into two halves results from the presence of the large auriculo-ventricular valves posteriorly and the large vessels anteriorly. The fibers comprising these bundles are clearly differentiated from the myocardium by their histological resemblance to the auriculo-ventricular bundle of the mammalian heart and by the presence of a connective tissue envelope. Further development leads to a greater and greater concentration of the fibers anteriorly with a disappearance of other parts of the ring, until in the mammalia the sole connection between auricles and ventricles is by a single collection of fibers originating in the auriculo-ventricular node and dividing at the anterior part of the base of the interventricular septum to form branches to each ventricle, the auriculo-ventricular bundle of His.

4. A sino-auricular junction, containing nodal tissue similar in structure to the sino-auricular and auriculo-ventricular nodes of the mammalian heart, is found at the right auricular border of the sinus venosus in the tortoise and selachian heart. In the eel, the nodal tissue is quite diffuse throughout the wall of the sinus, but accumulated somewhat more abundantly at the junction between the sinus and the

auricle. In the tortoise, the nodal tissue forms an almost complete circle at this junction. In the bird's heart, histological differentiation has proven of some difficulty, and it was at first thought that the system of nodal tissue was absent (19), (20). Subsequent work, however, revealed the presence of nodal tissue, although there seems to be considerable doubt as to whether a sino-auricular node is present separate from the auricular portion of the auriculo-ventricular node (atrial node) (23). In the lower mammals, nodal tissue is extensively distributed along the original sino-auricular junction, while in the higher mammals it tends to become concentrated near the termination of the superior vena cava.

5. Other remains of the sinus venosus, besides that forming the sino-auricular and a part of the auriculo-ventricular nodes, probably exist in the adult mammalian heart. The cross striated fibers at the mouth of the superior vena cava have been especially noted in this connection.

6. While Keith and Flack, and Koch have described rather extensive connections between the sino-auricular node and the auricular musculature, most investigators have been unable to trace any connection between the sino-auricular and auriculo-ventricular nodes, which shows any anatomical differentiation from the usual cardiac muscle. On all sides, short ramifications from the node disappear into the surrounding muscle. The supraventricular connections of the auriculo-ventricular node likewise appear to consist of numerous short branches extending out into the surrounding muscle. Curran (25), however, has described rather extensive bundle connections of the auriculo-ventricular node with the superior vena cava, right auricle and left auricle, and Thorel (26) found numerous branches from each node, consisting in large part of Purkinje cells, and making connection in the region of the coronary sinus. Thorel's findings have been denied by Aschoff and Monckeburg (27), Koch (15) and others.

THE ORIGIN OF THE CARDIAC IMPULSE IN THE LOWER VERTEBRATES

All of the older workers placed the origin of the beat in the tortoise and frog within the sinus venosus (28), (29) or at the junction of this chamber with the veins (30). More recently Garrey (31), working with the hearts of large tortoises and employing inspection and records by suspension, concluded that the beat began not in the sinus venosus, but in the right vein. This structure, according to Garrey, beats an appreciable interval before the sinus, and there may be recognized a veno-sinal interval comparable to the sino-auricular and auriculo-ven-

tricular intervals. The tortoise heart is thus regarded as consisting of four, rather than three, functional chambers, each beating alone and followed by the contraction of the next lower chamber after an appreciable pause. Meek and Eyster (32), using the electrical method, located the seat of initial impulse formation in the sinus venosus, and in a later work (33), in which large turtles were used, the exact location was determined in the sino-auricular junction. This latter position as the seat of origin of the impulse in the tortoise heart has also been arrived at by Schlomovitz and Chase (34), as a result of experiments in which the pacemaker was located by the method of local heating and cooling.

In the bird's heart Flack (35) was unable to demonstrate any specific or noteworthy effect on electrical or mechanical stimulation of the junction of the venae cavae with the right auricle, and also found that cold applied to this region was less effective in slowing the whole heart than when applied to the right atrium. Mangold and Kato (36), however, localized the seat of impulse formation in the bird's heart, by the method of local cooling, at the junction of the right superior and inferior vena cava with the right auricle. Since histological observations carried out by others (18), (19) had failed to reveal nodal tissue in this region, Mangold and Kato were inclined to believe that the origin of the heart beat in birds was in the usual cardiac muscle and not associated with nodal tissue. Subsequently, however, Mangold (37) associated the origin of the beat with nodal tissue, but rather with the auricular portion of the auriculo-ventricular node (atrial node) than with the sino-auricular node. This conclusion was based on the view of Aschoff (23), that the sino-auricular node is not present in the bird's heart, and the fact that the removal of the influence of the normal seat of initiation by local cooling (38) results in a rhythm in which the ventricular beats precede the auricular beats, and in which disturbances of conduction occur, with a more rapid ventricular than auricular rate. This type of rhythm is similar to that obtained by Zahn (39) in mammals as a result of cooling the auricular portion of the auriculo-ventricular node, and quite different from the type of rhythm obtained by cooling the sino-auricular node in these animals (see page 18).[3] Mangold (38) also

[3] The rapid heart rate normally present in birds make the study of the origin and conduction of the beat difficult in these animals. Mangold states that it is frequently necessary artificially to slow the rate of the beat. This is undoubtedly a procedure associated with possibilities of great error, since many experiments indicate that it may result in removal of the seat of impulse formation from its normal position. This fact, combined with the uncertainty as to the exact character and relation of the nodal tissues in these animals, places the whole problem in a somewhat unsatisfactory state.

found that the region of the bird's heart that is susceptible to local cooling, showed initial negativity when compared with the right atrium (see page 20).

EARLY EXPERIMENTAL OBSERVATIONS ON THE MAMMALIAN HEART

The anatomical discoveries described in a preceding section have profoundly influenced the point of view of those investigators who have attempted to throw light upon the problems of the origin and conduction of the excitation in the mammalian heart. The earlier workers directed their attention particularly to the lower part of the superior vena cava or the intercaval regions, in the belief that these parts represent principal homologues of the primitive sinus venous, while since the description of the nodes and the ventricular conduction system, the main question has been the importance of these structures for impulse initiation and conduction. It would seem advisable, therefore, to describe first certain of the more recent researches which were carried out before these structures were described or without special reference to them, linking them up in certain instances with a few more recent repetitions or extensions of earlier observations.

One of the first investigators to apply modern experimental methods to the study of this problem in the mammalian heart was MacWilliam (40). Employing the principle elucidated by Gaskell (41) in work on the cold-blooded heart, that the rate of beat of the whole heart was affected by heat and cold only when applied to the region initiating the impulse, he found that mild heat applied to the terminal portions of the vena cava caused acceleration of the whole heart, while it was ineffective when applied to the ventricles. He did not state whether or not other regions of the veins or auricles were tested. Adam (42) localized more accurately the area affected by heat, finding this within the intercaval region and the point of maximum sensitiveness near the inferior vena cava. Langendorff and Lehmann (43) excised the region described by Adam in artificially perfused rabbits' and cats' hearts, and found permanent auricular stoppage with temporary cessation of ventricular activity. If the hearts were perfused with defibrinated blood instead of saline solution, the results were less pronounced, slowing occurring in rabbits', but in only one of four cats' hearts, and in no case did ventricular quiescence result. In two rabbits, the region was excised from the heart in situ. In one of these the auricles stopped for over a minute, in the other the rate was much slowed. After the excision the ventricle showed a shortened compensatory pause following an

extrasystole, while previously ventricular extrasystoles had been associated with a full compensatory pause. This indicated that the ventricle assumed the function of impulse initiation, and further support for this view was found in the reaction of the ventricle to heat and cold after the excision in the intercaval region. Langendorff and Lehmann were of the opinion that neither the auricles nor the venae cavae have the power of independent automaticity, and in subsequent experiments of Lehmann, reported by Langendorff (44), it was found that the auricular appendages of both right and left auricles were not automatic, although automaticity could be developed within them by various external stimuli. Erlanger and Blackmann (45) repeated these experiments and obtained transitory stoppage of the ventricles with recovery at a slower rhythm. They also investigated the automaticity of various portions of the supraventricular regions of the heart by determining the rate of beat of isolated strips. The right atrium, near its junction with the superior vena cava, was found to have the highest automaticity in the sense that it contracted at a higher rate than other regions when isolated. The next most automatic region was found in the tissue near the coronary sinus. All parts of the right atrium, however, showed evidence of automaticity. Strips prepared from the left atrium rarely showed independent rhythmic beating. Erlanger and Blackmann noted that the auricles and ventricles often beat simultaneously after removal of the region of the mouth of the superior vena cava (auriculoventricular or "nodal rhythm"). Subsequent experiments by Erlanger (46), carried out with reference to the importance of the nodal tissue, are summarized on page 19. Hering (47) reported that a cut about 2 cm. long in the long axis of the superior vena cava, extending across the sulcus terminalis into the auricle, caused prolonged stoppage of the supraventricular regions in a single experiment on an excised, artificially perfused dog's heart. He obtained similar results in other experiments from single cuts made along the line of the sulcus terminalis.

A number of the earlier workers attempted localization of the seat of impulse initiation by observing that region which continued beating longest in the dying heart, on the assumption that this part was normally the most automatic. The results, as might be expected, were quite divergent. The "ultimum moriens" was placed by MacWilliam (40) and by Hering (48) at the junction of the superior vena cava and right auricle, and by Fredericq (49) in a region lower down between the two venae cavae. Hering (50), subsequent to the description of the sino-auricular and auriculo-ventricular nodes, located the region of the

last surviving contractions in most cases along the sulcus terminalis, in the remaining instances in the region of the coronary sinus (coronary sinus or atrial portion of the auriculo-ventricular node). Erlanger and Blackmann (45) made no more exact localization of the ultimum moriens than the region of the great veins as a whole. Hirschfelder and Eyster (51) observed contractions of the mouths of the great veins rarely in dying hearts, and when such contractions occurred there was usually complete dissociation between these and the contractions of the auricle. Koch (52) found that the region of the coronary sinus was the last to show contraction in the dying hearts of stillborn infants. In three such cases he cut away successively portions of the caval and intercaval regions, right atrium and interauricular septum. No change in ventricular rate resulted until the lower third of the septum and the region of the coronary sinus was removed, when ventricular activity ceased. In the light of subsequent work it seems probable that these hearts were in auriculo-ventricular or "nodal rhythm" when the observations were made. More recent experiments of Eyster and Meek (53), in which the seat of the ultimum moriens was compared with the seat of impulse initiation as determined by the electrical method, have proved that the ultimum moriens is variable in position and may not coincide with the region which acts as the pacemaker in the normal heart nor even in the dying heart. The somewhat gratuitous assumption of the association of these two is therefore not supported by direct observation. These experiments moreover showed that impulse initiation may occur in a region showing no visible contraction.[4]

Several other experimental observations, either preceding the discovery of the sino-auricular and auriculo-ventricular nodes or made without especial reference to these structures, are worthy of record. Fredericq (54) studied conduction between the two auricles and between the auricles and ventricles by partially separating these chambers from the interauricular septum in dogs' hearts under isolation and artificial perfusion. It was found that after cutting away the venae cavae and a large part of the right auricle bordering the veins that the two auricles contracted synchronously. Complete separation of the two auricles, by a cut along the interauricular septum, produced allorrhythmia, and the ventricles continued to beat in rhythm with the auricle that retained

[4] This observation suggests therefore that the differentiation of automaticity and contractility may be complete in certain parts of the mammalian heart, and that regions in which the former property is most highly developed may have lost the power of contraction.

connection with the septum. Only a small bridge of connecting tissue (about 1 cm. in breadth), and indifferently situated, was sufficient to maintain a common rhythm of the two chambers. These observations were confirmed by Hering (55). Similar observations carried out by Fredericq (56) more recently, in which dogs' hearts in situ were used and functional separation obtained by clamping, also indicated that complete separation of the right auricle from the interauricular septum abolished its influence on the left auricle and ventricles, which then beat independently and usually simultaneously. Separation of the left auricle from the septum usually resulted in quiescence of this chamber.

While it has been generally recognized that graphic records secured by suspension methods are difficult of application to the mammalian heart in any attempt to analyze the contraction wave within individual chambers, due to the rapid contraction, blood movements and periods of common activity in one or more chambers, there have been a few efforts to analyze curves of this type. Rehfish (57) found in graphic curves taken by suspension from the auricle, a small wave, particularly pronounced during the slowing obtained by vagus stimulation, which immediately preceded the wave due to contraction of the main mass of the auricle. He referred this wave to contraction of the region initiating the impulse, the mouth of the superior vena cava. A wave on the venous pulse described by Eyster (58), which preceded the auricular wave in slow-beating human hearts, had apparently a diastolic rather than systolic origin. Extrasystoles produced by stimulating of the veno-auricular junction or the mouth of the superior vena cava in dogs, produced the usual type of venous pulse without a wave preceding the auricular wave. A splitting of the auricular wave of the electrocardiogram, interpreted as indicating separate activity of the seat of impulse initiation and the main musculature of the right auricle, has been noted by Hering as a result of vagus stimulation and by Eyster and Meek (59) after partial isolation of the sino-auricular node.[5]

One of the most important observations contributing to our present interpretation of the heart beat, is that after destruction or suppression of the normal pacemaker, the function of impulse initiation in most

[5] Axial electrocardiograms, made with a very sensitive string galvanometer from the tortoise, show a wave preceding the auricular or P wave, due to the contraction of the sinus venosus (59). The fact that such differentiation of activity is not evident in the mammalian electrocardiogram is probably due to the more rapid spread of the impulse and a shorter interval comparable to the sino-auricular interval of the lower vertebrates.

cases is assumed by some region lying near the auriculo-ventricular boundary. The impulses originating from this region reach the auricles and ventricles at or near the same time, so that these chambers, instead of contracting in their usual sequence, beat exactly or nearly simultaneously. This rhythm is known as auriculo-ventricular or nodal rhythm, and was first clearly described by Englemann (60). After the application of the first Stannius ligature (ligature between sinus venosus and auricle) in frogs' hearts, he found that in certain experiments the auricles and ventricles beat simultaneously or nearly so. On the assumption that the excitation spread from its origin in all directions at the same rate, Englemann concluded that it must arise under these conditions on the ventricular side of the auriculo-ventricular ring. A year later, Lohmann (61) observed that spontaneous beats that broke through vagal inhibition, both in the tortoise and rabbit heart, often showed a greatly reduced As-Vs interval. He concluded that the seat of impulse formation in these was in the "Brückenfasern" connecting auricles and ventricles and stated that accordingly as the ventricular, middle or auricular portion of these structures became automatic, the intervals were negative (i.e., Vs-As), zero, or positive in value. By stimulating the auriculo-ventricular junction with electrical currents, he produced a rhythm in which the intervals were permanently shortened. A clinical case in which the a-c interval of the venous pulse was zero was reported by Mackenzie (62) in the same year, and Hering (63) in 1905 reported shortened As-Vs intervals on accelerator and vagus stimulation in dogs. Lohmann (64), in a later work, secured auriculo-ventricular rhythm in rabbits' hearts as a result of the application of formalin to the mouth of the superior vena cava. He made no reference in this paper either to the sino-auricular or auriculo-ventricular nodes, but seems to have been the first to recognize clearly that the destruction of the influence of the normal pacemaker results in the removal of the function of impulse initiation to some region between the auricles and ventricles. He stated that this may occur either as a result of the suppression of the influence of the normal pacemaker, or by heightening the automaticity of the lower region by electrical stimulation of the auriculo-ventricular boundary. Further observations on the seat of impulse formation in auriculo-ventricular rhythm are summarized in the next section.

OBSERVATIONS ON THE IMPORTANCE OF THE SINO-AURICULAR AND AURICULO-VENTRICULAR NODES

With the description of the nodes and the auricular and ventricular connections of the auriculo-ventricular bundle, renewed interest was aroused in the problem under consideration. This increased interest is indicated by the large number of recorded observations and the application of new methods or the elaboration of older methods of study. The methods of inspection, mechanical recording, the effect of heat and cold, the study of automaticity in isolated strips, isolation of parts in the excised heart or the heart in situ by cutting or clamping, have been employed in more extensive observations and with greater refinement. The most important of the new methods is that which determines the region of initiation and the path of conduction of the excitation by recording the accompanying electrical current. Although action currents from the heart had been recorded with the capillary electrometer as early as 1879 by Burdon-Sanderson and Page (65), and later by Waller (66) and by Bayliss and Starling (67), it was not until the modification and adaptation of the string galvanometer to physiological purposes by Einthoven, that this method was applied intensively to a study of the origin and conduction of the impulse. The electrical method has been applied in two ways. In one the propagation of the wave of negative potential has been studied by the application of electrodes directly to the heart. In this method the onset and direction of the wave determines the point of initial negativity, and the path of progression and the velocity can be calculated by the time required for the wave to pass between two or more electrodes. This method had been widely used previously in skeletal muscle and in nerve. In the other, leads from the body surface are used to record the summated action currents of the heart in the form of the electrocardiogram, and the curves then submitted to analysis. In general it may be said that the former has yielded more definite information, while the latter has been of most value in adding further evidence, and information relative to the path of conduction in the human heart.

While we do not wish to minimize the importance of many of the results obtained by other methods of study, we believe that the most extensive and reliable information that we possess at present has been gained through the adaptation of the electrical method. We shall therefore devote a separate section to this subject and consider here only the more recent observations made by the application of other methods.

Tawara (68), following his discovery of the auriculo-ventricular node, suggested that it represented the primary motor center of the heart, and Keith and Flack (7) ascribed the same function to the sino-auricular node in their paper first describing this structure. Several years elapsed, however, before experimental observations were carried out. In 1910 a number of papers dealing with experimental studies appeared almost simultaneously. The first of these to be reported were the experiments of Jaeger (69), who destroyed a wide region, including the sino-auricular node, by application of the thermo-cautery in a series of dogs and cats. From the diagrams given in his paper, it is evident that the region cauterized included the lower portion of the lateral surface of the superior vena cava, a considerable portion of the base of the right atrium and auricular appendage and the upper two-thirds of the intercaval region. Perforation and rapid death from hemorrhage occurred in all but a few experiments. No disturbance of rhythm occurred in any case. While Jaeger's experiments were directed toward the question of the importance of the sino-auricular node and his conclusion was to ascribe no important function to it, it should be pointed out that from his results the same conclusion would necessarily apply, not only to the node, but to a wide area including the mouth of the superior vena cava, the intercaval region and a considerable portion of the caval border of the right atrium and appendage. Flack (70), one of the discoverers of the sino-auricular node, reported at nearly the same time[6] a long series of experiments in which a study was made of the effects of the application of cold, of electrical stimuli, of atropin and muscarin and of clamping and ligaturing the node, and the influence of these procedures on the control of the extrinsic cardiac nerves. The experiments were carried out on the hearts of dogs, cats, rabbits and fowls in situ. The influence of cold in decreasing the rate of beat of the whole heart was confined to the sino-auricular node. Electrical stimulation caused inhibition or acceleration, depending on the strength. The same strength of stimulation applied to neighboring regions had little or no effect. Little effect was noted on ligaturing or clamping the node. Unfortunately Flack's records were made at too slow a speed to detect any changes in the rate of auriculo-ventricular conduction or the auriculo-ventricular sequence. His conclusion was that normally the dominating rhythm of the heart arises in the node. Flack (71) continued his studies in a series of papers published a year later, and concluded finally, that while the

[6] The first electrical studies were made almost contemporaneously, as will be described later.

sino-auricular node may possess a high degree of automaticity, there are other parts in which this property is well developed, and that its main rôle is that of a "cardio-regulating center" rather than a cardio-motor center. Flack thought that it is through this region that the dominant rhythm of the heart is most readily influenced, this influence being exerted mainly by the intimate association of the node with the extrinsic cardiac nerves as evidenced by its neuromuscular structure. He stated that since its excision produced an insignificant physiological effect, it was probably of no clinical importance. A similar view had been expressed by Ivy Mackenzie (72) on the basis of the histological structure of the node. Mackenzie suggested that its function was to act as a receptor for sensory impulses necessary for coördinated heart activity, in a way comparable to the muscle spindles of Kühne and Sherrington in skeletal muscle.

Hering (73) attempted to destroy the sino-auricular node by burning with a hot wire. The experiments were made on dogs' hearts in situ. He found that there was little change except in the interval between auricular and ventricular contraction. This interval was usually reduced by injury to the node. The reduction was not infrequently prededed by an initial transitory lengthening. Sometimes the auricles and ventricles beat simultaneously or nearly so, or the ventricular beat preceded the auricular beat (Vs-As sequence). He suggested that in this rhythm (auriculo-ventricular), the seat of impulse formation was in the auriculo-ventricular node, and on the basis of his previous observations that the main delay in the passage of the impulse between auricle and ventricles occurs in the auriculo-ventricular node,[7] suggested that the different As-Vs intervals may be accounted for by differences in the exact location of the seat of impulse initiation within this structure. If this is in the upper part of the node, an As-Vs interval approaching

[7] Shortly after the appearance of Tawara's description of the auriculo-ventricular node, Hering (74) suggested that the greater part of the delay in the passage of the impulse from the auricle to the ventricle probably occurred in this node, and in 1910 (75) he offered experimental support for this view. The isolated perfused hearts of large dogs were opened and stimuli applied directly to the region of the auriculo-ventricular node and the time from the incidence of the stimulus to contraction of the ventricle recorded. If the bundle were cut near its origin from the node, stimulation of the ventricular end was followed after a very short interval by ventricular contraction, while the corresponding interval between the stimulus and auricular contraction was much larger when the auricular end of the cut bundle was stimulated, due to the fact that in the later case the impulse had to pass through the node. The difference amounted to four or five times.

the normal might occur, while if in the mid-region one would expect the interval to approach zero, and finally if in the lower portions of the node, a Vs-As sequence would be the probable result. This view was merely an adaptation to the nodal structure of the previous suggestion of Lohmann. Histological examination of these hearts was made by Koch, who found that while considerable damage to the sino-auricular node was produced, it was not entirely destroyed in any experiment. As a result of these experiments, Hering inclined to the belief that the sino-auricular node is the normal seat of impulse formation, although he suggests that possibly it serves merely to conduct and distribute the impulse which actually arises in the superior vena cava.

Magnus-Alsleben (76), as a result of experiments on isolated rabbits' hearts perfused by Langendorff's method with diluted ox blood, reported negative results from excision of the sino-auricular node and other regions of the auricles and venae cavae. The As-Vs intervals were unchanged, although the rate of beat was usually slowed. In one series of experiments, both auricles were cut away piece by piece without any constant or noteworthy influence on ventricular activity. The general conclusion drawn was that there is no single area in the veno-auricular region of predominant importance for impulse initiation. While admitting that probably the normal seat of origin of the impulse may be at the veno-auricular junction, he concluded that when this is removed the seat of automaticity passes downward by steps ("schrittweise") without disturbance of the beat, although the rate of contraction is usually somewhat slower when initiated from the lower regions.

Cohn, Kessel and Mason (77), (78) performed an extensive series of experiments on the effects of removal of the sino-auricular node in dogs' hearts isolated and under artificial perfusion with salt solution. A total of thirty experiments as carried out in which attempts were made to excise the sino-auricular node with as little damage as possible to surrounding tissue. Diagrams are given to show the region excised and subsequent histological examinations were made. Transitory quiescence of the whole heart, varying between four seconds and three minutes, and usually less than thirty seconds, occurred in twenty-four of the thirty experiments. The period of quiescence was interpreted as representing the time required for the auricle or other tissue to take up the pacemaking function. Simple incision through or across the node was without effect. Unfortunately, statements of the normal As-Vs interval and its change as a result of excision of the node, are given in only a few instances, although it is stated that in fourteen

experiments As preceded Vs, in seven Vs preceded As, and in two the auricles and ventricles contracted simultaneously. It seems certain therefore that auriculo-ventricular rhythm occurred in a number of these experiments and its occurrence is excluded in none. The conclusion of these workers that no special region takes up the function of impulse initiation after removal of the sino-auricular node, is therefore not justified in the light of preceding and subsequent experimental data. The rate of beat following the excision of the node was permanently reduced. Excision of neighboring tissue failed to cause similar results to that following excision of the node. Similar experiments were performed by Moorehouse (79), who found that excision of the sino-auricular node produced at times transitory quiescence of the auricles, at other times slowing without quiescence and in some instances no effect whatever upon auricular rate. He concluded that the slowing or quiescence, when it occurs, may be due to disturbance of the blood supply or to mechanical stimulation, and does not indicate any important physiological significance of the sino-auricular node. The As-Vs sequence and intervals were also not reported in these experiments. Removal of areas of tissue outside the node gave the same inconstant results.

The influence of cold applied locally to the sino-auricular node has been studied in detail by Brandenburg and Hoffmann and by Ganter and Zahn. Brandenburg and Hoffmann (80) worked on isolated rabbits' and dogs' hearts under artificial perfusion. Cold was effective in reducing the rate of the whole heart only when applied to the region corresponding to the sino-auricular node. When marked cooling was carried out, auricles and ventricles beat simultaneously, in which case the authors assumed the impulse was arising within the auriculo-ventricular node. The transition from sino-auricular to auriculo-ventricular rhythm was abrupt, and associated with a reduction of the rate to about one-half the normal. The change to auriculo-ventricular rhythm was referred to depression of automaticity in the sino-auricular node, as a result of the cooling, and the assumption of the function of impulse initiation by the region of lower automaticity. On cessation of cooling the normal rhythm recurred. Failure to obtain this result, they state, may be due to the heart not being in normal rhythm at the beginning of the experiment, or because of abnormal irritation of the right auricle from cutting or crushing, the seat of impulse formation on cessation of automatic activity of the sino-auricular node may pass to some region of the auricle rather than to the auriculo-ventricular node.

When auriculo-ventricular rhythm was established as a result of cold applied to the sino-auricular node, they found that it could be abolished, without however a return of sino-auricular rhythm, by clamping, scraping or otherwise submitting the neighboring auricular tissue to irritation. The As-Vs interval returned to almost normal value. Under these conditions they were unable to determine any localized region which, by the application of cold, could be proved to be controlling impulse formation. The seat of impulse initiation was apparently widespread, involving the whole of the right and possibly the left auricle. They were thus inclined to believe that while under approximately normal conditions only the two nodes manifested automatic power, that other regions (right and possibly left auricle) could develop automaticity when exposed to abnormal influences of a nature tending to heighten tissue irritability. Brandenburg and Hoffmann also studied the effect of partial isolation of the sino-auricular node by a series of cuts along the auricular border. After each cut determination was made as to whether or not the node was still acting as the cardiac pacemaker by the application of cold. It was found that only a small bridge of connecting tissue was necessary to maintain functional integrity. The experiments of Ganter and Zahn (81) on the hearts of cats, rabbits, goats, dogs and apes in situ, agree in many details with those of Brandenburg and Hoffmann on isolated hearts. Ganter and Zahn found that the region normally susceptible to cooling corresponds to the area including the sino-auricular node, as determined by histological examination of the hearts experimented upon. As a result of cold applied to this region, a new type of rhythm develops, characterized by marked reduction or disappearance of the As-Vs interval. The upper portion, or "head" of the node, was found the most susceptible region to relatively slight cold, and when the rhythm was normal, cooling the lower end of the node alone produced no change in rate. The transition from sino-auricular to frank auriculo-ventricular rhythm was found to be abrupt or gradual, depending upon the rapidity of cooling of the sino-auricular node. If the cooling were gradual, the change in As-Vs interval was not abrupt, as found by Brandenburg and Hoffmann, but occurred through several stages. When auriculo-ventricular rhythm was established, as shown by great reduction in the As-Vs interval, warming or cooling the region of the auriculo-ventricular node, by means of a specially designed thermode introduced through a hole in the auricular appendage and applied to the base of the inter-auricular septum, was effective in influencing the rate of the whole heart. While, as we have

seen, a number of workers had suggested that the auriculo-ventricular node was the seat of impulse initiation in auriculo-ventricular rhythm, this was the first direct experimental evidence to support this assumption. When the pacemaker resided in the sino-auricular node, as normally occurs, cooling the auriculo-ventricular node caused nothing more than disturbances in auriculo-ventricular conduction. The final conclusion was that the capacity of rhythmic impulse formation is at least in high degree associated with the sino-auricular and auriculo-ventricular nodes. These experiments were later repeated and extended by Zahn (39). Using the warming and cooling apparatus employed in the former work combined with an endoscope, he was able to apply heat or cold to different parts of the auriculo-ventricular node and neighboring regions in the heart in situ. A few experiments were also carried out on isolated hearts. The sino-auricular node was first destroyed by the application of a hot glass rod or of formalin, or in isolated hearts it was excised. The criterion of its destruction in the heart in situ was the absence of the usual effects of the application of heat and cold. Zahn was especially concerned with determining the seat of impulse formation in those cases where, after destruction of the sino-auricular node, the As-Vs interval was of approximately normal length. Brandenburg and Hoffmann, as noted above, believed that under these conditions the impulse arose in a widespread region of auricular tissue. Zahn found, as did these workers, that destruction of the impulse forming power of the sino-auricular node did not always produce a frank auriculo-ventricular rhythm, or if this was present at first, the As-Vs interval frequently gradually increased until it approached the normal. He found, however, that in all such cases a localized area could be determined which dominates the rhythm. This region is the tissue surrounding the mouth of the coronary vein. Histological examination of the hearts from these experiments was made by Koch, who found that this region contained abundant nodal tissue, and identified it with the coronary sinus portion of the auriculo-ventricular node as previously described by Aschoff (22).[8] Zahn concluded that after elimination of the sino-auricular node some part of the auriculo-ventricular node assumes the function of impulse initiation, and the As-Vs interval in the rhythm which results depends upon the portion of the latter node that is active. After removal of the sino-auricular node, warming the coronary sinus portion of the auriculo-ventricular node produced a rhythm in which

[8] For a detailed description of the atrial node and the relation of nodal tissue to impulse initiation, reference may be made to Koch (16).

the As-Vs interval was only slightly less than normal, while warming the middle and lower regions produced a much shortened As-Vs interval, the disappearance of this interval or the occurrence of a Vs-As rhythm. All parts of the auriculo-ventricular node thus seem to have the power of automaticity without great variation in degree in different parts.[9]

Further contributing evidence pointing to the sino-auricular node as the normal seat of impulse initiation has been furnished by Sansum. It was shown by Englemann (83) in 1897 that in the frog premature beats produced by stimuli applied to the auricles and ventricles were followed by a pause sufficiently longer than the usual diastolic pause to restore the original rhythm (compensatory pause), while if such extra stimuli were applied to the chamber initiating the rhythm (the sinus venosus), the pause following the extra beat was little if any larger than the usual diastolic pause. Sansum (84) attempted to apply this same principle to the mammalian heart, and while it was found that all auricular extrasystoles showed less than a full compensatory pause, this pause was shortest in extra systoles arising from stimuli applied to the sino-auricular node.

The line of work which has yielded results most opposed to the view that the sino-auricular node is to be regarded as the most automatic region of the heart and the normal seat of impulse initiation, is the study of relative automaticity in different regions of the heart by isolation of strips and determination of their rate of contraction when thus isolated. The early work of Erlanger and Blackmann (45), made without reference to the sino-auricular node, has already been quoted (page 8). These experiments were repeated and extended by Erlanger in 1910 (46), in a study of strips isolated from thirty-nine cats' hearts. He found that strips from the intercaval region, mouth of superior vena cava, right atrium, inferior vena cava, coronary sinus and interauricular septum showed spontaneous contractions when isolated. Strips from the right auricular appendage (except in one instance), the left auricle and the pulmonary veins were inactive. In the rhythmic strips the rate of beat varied greatly, but was approximately the same in

[9] A critical consideration of the method of the application of heat in locating automatic regions is given by Schlomovitz and Chase (82). These authors also confirm the location of the normal pacemaker in the sino-auricular node, but find that any of the "sinus tissue" of the veno-auricular regions may respond to higher temperatures, presumably due to an increase in its automaticity as a result of the application of heat. Another explanation seems possible, however, namely, conduction of the heat to the neighboring sino-auricular node.

those from the superior and inferior venae cavae and the coronary sinus region. Stimulation of quiescent strips from these regions would usually start rhythmic beating, which continued after the cessation of the stimulus, and it was found that the most constant result in this respect was obtained from strips of intercaval tissue from a region parallel with and on the venous side of the sulcus terminalis. A repetition of this work with histological control was made by Moorehouse in 1912 (79). Strips were made from the isolated dog's heart and the rate of contraction and response to rhythmic stimulation determined. Strips containing the sino-auricular node showed no higher rate of contraction than those made from neighboring tissue in which no nodal tissue was present. In a subsequent paper Moorehouse (85) studied the influence of various pharmacological agents on similar strips and could find no difference in behavior of those containing and those not containing nodal tissue.

THE ORIGIN AND CONDUCTION OF THE IMPULSE IN MAMMALIAN HEART AS STUDIED BY THE ELECTRICAL METHOD

Attempts to locate the primary seat of impulse initiation in the mammalian heart by galvanometric methods were made at almost the same time by Wybauw (86) and by Lewis, Oppenheimer and Oppenheimer (87).[10] These determinations were based on the assumption that a condition of electrical negative potential is associated with activity, and that when the activity is propagated in the form of a wave, it is accompanied by this alteration in electrical condition.[11] The problem

[10] The use of the method was apparently independently applied by Wybauw and Lewis. Statements in each paper indicate that each worker was aware of the investigations of the other. Lewis used the method in a publication (88) preceding Wybauw's preliminary note, in which he found that in two instances after recovery of normal rhythm following experimental auricular fibrillation in dogs, a point on the superior vena cava was negative before a point on the inferior vena cava. In the paper on the electrocardiograms of auricular extrasystoles, Lewis (91) states that preliminary observations had been undertaken with this method, but hearing that Wybauw's observations were nearing completion, this line of investigation was not pursued. The proof that initial negativity is associated with the region of the sino-auricular node, was first furnished by Wybauw. In the paper of Lewis, Oppenheimer and Oppenheimer, reference is made to the preliminary publication of Wybauw.

[11] This association had been long established in skeletal muscle and nerve by Du Bois Reymond (1873) and subsequent workers, and Kölliker and Müller (89)

in the heart was therefore to find the region showing initial negativity at the onset of the beat, the location of which should be identical with the region manifesting initial activity. Electrodes were placed at various points on the right auricle, superior and inferior venae cavae and intercaval region, and connected in pairs with a string galvanometer. The point showing initial negativity was determined by the direction of the initial movement of the string in the curve obtained, it having been determined beforehand the direction of movement when one pole of the galvanometer was made electro-negative to the other. Wybauw found in thirty out of thirty-eight determinations in the dog's heart in situ, that an electrode on the sulcus terminalis, at a point corresponding to about the middle of the sino-auricular node, showed initial negativity when compared with other regions. He found that the wave of negativity spread out rapidly from this region and reached the right atrium within a few hundredths of a second, indicating the absence of a sino-auricular interval comparable to that present in the lower vertebrates. Lewis, Oppenheimer and Oppenheimer determined the seat of initial negativity in the hearts of seven dogs in situ. Seven points along the sulcus terminalis, six points along the intercaval region, four points on the right atrium and two points on the pulmonary vein were investigated. At the completion of each experiment the hearts were examined histologically in regard to the position of the electrodes in relation to the sino-auricular node. The upper end of the sino-auricular node, near its widest region, was found negative to all other points in six or seven experiments. In one experiment a point near the mouth of the pulmonary vein was negative when compared with the sino-auricular node. These workers, therefore, placed the seat of normal impulse initiation in the dog in the upper part or cephalic end of the sino-auricular node, and not in the mid-region, as had been found by Wybauw.

showed that the heart also developed action currents during normal contractions. The constant association of tissue activity with an "action current" seems to have been firmly established and is a generally accepted phenomenon. Direct evidence of the association between these two phenomena in the sino-auricular region of the mammalian heart has been recently obtained by Lewis (90). The auricular contraction wave, as recorded by the approximation of a grid of fine hairs or threads attached to points on the heart surface, was compared with the spread of the wave of negativity. While considerable difficulty attaches to the interpretation of graphic curves of this type when recording movements of small masses of muscle attached to larger moving masses, evidence was obtained that the two phenomena are directly associated, the contraction wave following the negative variation at any point by approximately 0.02 second.

In a preceding publication Lewis (91) had attacked the problem in another way. Axial electrocardiograms from dogs were recorded and extra or premature contractions of the heart produced by stimuli applied to different regions of the auricle and venae cavae. It was found that the auricular complex (P wave) resembled that obtained during the normal beats only in extrasystoles which were induced by stimulation of the region of the sino-auricular node. It was concluded therefore that the normal impulse arises in this region.

The relation of the region of the sino-auricular node to initial negativity was further confirmed by Eyster and Meek (92). In thirty-eight experiments on dogs' hearts in situ, some part of the sino-auricular node was found to precede one or more regions lying outside in all cases. Determinations made on the region of the node and its immediate neighborhood, controlled by histological examination, placed the initial change in that part of the node containing the greatest mass of tissue, usually the upper part. Initial negativity of the sino-auricular node, when compared with other regions, was also found when the galvanometer was made unusually sensitive (one millivolt giving a deflection of 1050 millimeters). This observation was believed to meet an objection to previous experiments raised by Erlanger (93), namely, that the excitation may start from outside the node and become concentrated within it, becoming of sufficient strength to influence the galvanometer only when such concentration had occurred. The endocardial surface of the node was also found initially negative to epicardial points outside the node. Experiments were also carried out by Eyster and Meek (59), (92) in an attempt to determine the rate and paths of conduction from the sino-auricular node to other regions, particularly the right auricle and auriculo-ventricular node. In addition to comparisons of negativity, the incidence of the main electrical change at various electrodes compared with the onset of mechanical systole of the right auricle recorded by suspension was used as a basis for calculating the rate of propagation of the electrical wave It was found that while the spread from the sino-auricular node to surrounding structures was of a radial nature and that no restricted path or paths of conduction could be made out, there were apparent certain differences in rate which were regarded as significant. The wave of negativity was found to be propagated throughout the sino-auricular node and to contiguous venous regions with considerable rapidity, reaching the mouth of the superior vena cava within 0.01 second, while its spread to the auricular muscle was somewhat slower, a sino-auricular interval averaging 0.027 second being

determined in fourteen experiments.[12] This delay in passage to the auricle was such that in the majority of the experiments the auriculo-ventricular node became involved before the right atrium. It was concluded, therefore, that the excitation spreads by independent paths from the sino-auricular node to the right atrium and to the auriculo-ventricular node, and does not pass through the atrium to the ventricle as ordinarily stated.

Lewis, Meakins and White (95) carried out a series of experiments on dogs' hearts in situ, in which the spread of the wave of negativity from the sino-auricular node was studied by determination of the direction of movement of the galvanometer, and the rate estimated by comparison of the incidence of the wave at given points with the summit of the P wave of the axial electrocardiogram. They were unable to determine any essential difference in the rate of propagation such as had been described by Eyster and Meek,[13] and concluded that the excitation spread in all directions from the sino-auricular node at approximately the same velocity. The actual average rates as determined show, however, the lowest rate of propagation to the superior vena cava (588 mm. per sec.), and the most rapid propagation to the intercaval region (1232 mm. per sec.) and to the auriculo-ventricular node (1101 mm. per sec.). The rate to the right auricle averaged 859 mm. per second.

Meek and Eyster (96) proved by the method of initial negativity that when the sino-auricular node was rendered functionless by the application of formalin, by crushing, cutting or clamping around the node or by vagus stimulation, the seat of impulse initiation invariably removed to the region of the auriculo-ventricular node. When the

[12] Garten (94) reported a similar determination by Sulzi, working under his direction on six dogs' hearts in situ. Comparing the negativity in the sino-auricular node with a point on the right atrium, 18 mm. distant, a difference of 0.022 to 0.026 second was found. Pairs of electrodes, the termination of which were placed in contact, were applied to the two regions and each pair connected to a galvanometer.

[13] There is little doubt that the comparison of the electrical change with other events in the auricular cycle in an attempt to measure velocity is subject to considerable error and can be taken as yielding only approximate results. This error arises mainly from two sources, the difficulty of exact location of changes in mechanical suspension curves or of the P wave of the electrocardiogram, and the fact that the relation between mechanical events, local action currents, and the electrocardiogram is subject to considerable variation. Measurement of velocity by using two galvanometers and determining the time of passage of the electrical change between pairs of electrodes connected to these, is far less open to criticism.

As-Vs interval approached the normal, the seat of initial negativity was found to be in the coronary sinus portion of the auriculo-ventricular node, while with great reduction in the As-Vs interval, it was located in the ventricular portion of this node. In no case, after the sino-auricular node had ceased to function, was it possible to locate initial negativity outside the area containing the auriculo-ventricular node. In another series of experiments (97), these investigators proved that the application of slight cold locally to the upper portion of the sino-auricular node, the normal seat of initial negativity, resulted in depression of automaticity in this region, and initial negativity now appeared in lower portions of the same node. This proof was furnished by the simultaneous use of two string galvanometers, one comparing the upper with the lower end of the sino-auricular node, the other comparing the sino-auricular and auriculo-ventricular nodes. As a result of cooling the upper end of the node, the direction of movement in the first galvanometer reversed, showing initial negativity in the lower part of the node. That the seat of initial activity had not, however, been removed from the sino-auricular to the auriculo-ventricular node, was proved by the fact that the direction of movement in the second galvanometer remained constant. This shift of the seat of initial activity from the upper to the lower part of the sino-auricular node was associated with lengthening of the cycle and shortening of the As-Vs interval. After the cooling was allowed to pass off, the original condition was resumed, indicating a return of the function of impulse initiation to the upper part of the node. It seemed probable, therefore, that all gradations of As-Vs interval from the normal to zero and reversal were accountable for on the basis of impulse initiation in some part of the sino-auricular or auriculo-ventricular nodes.

In a subsequent paper Eyster and Meek (98) carried out further studies in reference to conduction of the impulse from the sino-auricular node to the right atrium and auriculo-ventricular node. Ten experiments were made on dogs' hearts in situ, in which a series of points, as nearly equidistant as possible from the sino-auricular node and on the auricular and venous sides, were compared as to the reception of the excitatory process by the method of initial negativity. As a preliminary measure, the fact that a normal rhythm was present was determined by proving that the upper portion of the node was negative before all other points and before the auriculo-ventricular node. These experiments showed that points on the venous or intercaval side of the node usually manifest negativity before equidistant points on the auricular side,

and that the excitation should be regarded therefore as spreading more rapidly in this direction than to the auricle. There was to be found no restricted region surrounding the node which invariably received the excitation before other regions, indicating no localized, restricted path of conduction. In forty-five experiments, also on dogs' hearts in situ, the sino-auricular node and immediately contiguous tissue was partially or completely isolated by successive cuts or by clamping. In ten of these the occurrence of activity in the sino-auricular node, the superior vena cava, the right atrium, the coronary sinus region and the auriculo-ventricular node was determined by initial negativity. In the remaining thirty-five, the time for propagation of the excitation from the sino-auricular node to the right atrium and to the auriculo-ventricular node and from the right atrium to the auriculo-ventricular node, was determined by applying two electrodes, separated by several millimeters to each of these regions, and connecting each pair to separate galvanometers. The comparison of the time of occurrence of the excitation in any two regions was made from simultaneous curves of the two galvanometers. The data from the normal, before partial or complete isolation of the sino-auricular node, supported the view previously advanced, namely, that conduction to the right auricle is delayed and that the auriculo-ventricular node receives the excitation by a path separate from this chamber. The auriculo-ventricular node received the excitation before the right atrium in nine out of twelve experiments in which initial negativity was determined. In twenty-four experiments in which time intervals were measured, the right atrium, however, received the excitation first in fifteen, and the average of all placed this chamber 0.003 second before the auriculo-ventricular node. The average time for passage of the excitation from the sino-auricular to the auriculo-ventricular node was 0.023 second in twenty-eight experiments, and the corresponding time from the sino-auricular node to the right atrium was 0.020 second in nineteen experiments. Isolation of the node was carried out by making successive cuts or clamping the tissue on the four sides of the node, determinations being made after each procedure. The following is a summary of the more important results. A single cut, made along the intercaval border of the sulcus terminalis, produced permanent auriculo-ventricular rhythm in 25 per cent of the experiments and in 12.5 per cent more this rhythm was established temporarily. A single cut across the sulcus above the head of the node produced transitory auriculo-ventricular rhythm in 40 per cent, but in no case was it permanent. A single cut across the sulcus below the

node produced permanent auriculo-ventricular rhythm in 10 per cent and transitory rhythm of the same type in 10 per cent. A single cut along the auricular border of the node failed to produce either transitory or permanent auriculo-ventricular rhythm in any of seventeen experiments. This last cut, however, increased the time of conduction from the sino-auricular node to the right auricle, but did not affect the corresponding interval between the sino-auricular and the auriculo-ventricular nodes. The latter interval was lengthened, however, by the other cuts, particularly the one across the sulcus below the node. Isolation of the sino-auricular node on three sides, the right atrial connection remaining intact, caused a marked increase in the conduction time between the sino-auricular and auriculo-ventricular nodes. This procedure also produced auriculo-ventricular rhythm in 82 per cent. Interruptions of the connections of the sino-auricular node on all four sides produced auriculo-ventricular rhythm in all cases. From these results it was concluded that there is no single, well-defined and localized path which serves to the exclusion of others for the conduction of the excitation from the sino-auricular to the auriculo-ventricular nodes; at least it may be said that if there is such a path, there are other paths which are capable of conducting the excitation with almost equal facility. The experiments seem to establish, however, with equal clearness, that the spread of the excitation from the sino-auricular node to the right auricle and auriculo-ventricular node is by separate paths, and the former does not normally receive the impulse through the latter, but may do so (in approximately 27 per cent of all cases) when the normal path of the impulse is interrupted. The excitation, arising normally in the sino-auricular node, spreads directly to the right auricle, probably in a diffuse manner, from the auricular side of the node. It also spreads to the auriculo-ventricular node by paths, probably also diffuse in nature, connecting with the upper, intercaval and lower borders of the node.

In a subsequent communication, in which the same methods were employed, Eyster and Meek (99) furnished experimental proof of the occurrence of partial and complete sino-auricular and sino-ventricular block by independent registration of the activity of the sino-auricular node, auriculo-ventricular node and right auricle, by means of the action current developed in these regions.[14] Partial block was obtained in

[14] A number of examples of supposed sino-auricular block in mammalian experiments and in clinical cases have been reported—Erlanger and Blackmann (45), Eyster and Meek (100), Eyster and Evans (101) and Levine (102);—but in all these the condition has been inferred from dropped beats in an otherwise regular rhythm.

only one of twenty-one experiments. The probable factors which determine the comparative rarity of this condition in the mammalian heart were considered in detail. The infrequent occurrence of partial sino-auricular and sino-ventricular block was believed to be due to *a*, the diffuse nature of the conducting path between the sino-auricular and auriculo-ventricular nodes; *b*, the possibility in a certain percentage of cases, of conduction between these two nodes by way of the right atrium, when the normal sino-ventricular path is destroyed; and finally *c*, the close approach of the automaticity of the auriculo-ventricular node to that of the sino-auricular node. Complete sino-auricular and sino-ventricular block was regarded as a potential possibility in any instance of auriculo-ventricular rhythm, and its presence was proven in two experiments. In these the region of the sino-auricular node was found to be generating impulses at approximately its original rate and at a faster rate than the auriculo-ventricular node. The auricles and ventricles were following the rhythm of the latter node. It was also shown that the right auricle, while normally receiving its excitation directly from the sino-auricular node could receive the same from the sino-ventricular conduction path, but under these circumstances the period of conduction from the node to the auricle was much prolonged. These experiments were regarded as furnishing still further evidence of separate sino-auricular and sino-ventricular paths.[15]

Partial or complete removal of the sino-auricular node under aseptic precautions in dogs, with recovery and subsequent electrocardiographic observations, has been carried out by Eyster and Meek (104). The removal of a considerable portion of the node results in the temporary establishment of a rhythm in which auricles and ventricles beat simultaneously or nearly so. After a period, varying from several hours to several days, the interval lengthens, but seldom equals normal. In most instances subsequent examination by the method of initial negativity has shown that the permanent rhythm established had its origin in the coronary sinus part of the auriculo-ventricular node. In a few instances, however, the pacemaker resided permanently in the portion of the sino-auricular node which remained intact. Complete removal of the node is associated with technical difficulties, but has been successfully accomplished in several instances and a permanent auriculo-ventricular rhythm observed as long as six months. At the end of the period of observation in each experiment, the seat of impulse initiation

[15] A clinical case, probably to be interpreted as sino-ventricular block, has been described by Naish (103).

was determined in these animals by electrodes applied directly to the heart. Each experiment was also subjected to histological control. Control experiments, in which a portion of the right atrium near the node was removed, showed no change in rhythm from the normal. These experiments are interesting, not only in indicating the paramount importance of the node-bearing tissue, but also that only a relatively small portion of it is necessary to maintain normal rhythm. As with most physiological mechanisms, an excess beyond that necessary to maintain functional activity under usual conditions is present.

CONDUCTION IN THE VENTRICLE

Conduction of the excitation to the ventricles in the mammalian heart occurs, probably exclusively, by way of the auriculo-ventricular bundle.[16] Tawara showed that the branches of the bundle do not terminate, as had been supposed, near the auriculo-ventricular boundary, but continue as an extensively ramifying system making connection with the Purkinje network in the subendocardium. The initial branches pass out mainly to the papillary muscles. The branch to the right ventricle passes along the right side of the interventricular septum to the base of the chief papillary muscle. The left division breaks into two chief strands which pass to the anterior and posterior papillary muscles. Other smaller branches pass to the septum. The larger branches are enclosed by endocardial sheaths which may be brought out clearly by injection methods (108). The Purkinje reticulum or network lines nearly the whole subendocardial space, and occasionally strands run for several millimeters into the ventricular muscle.

In the lower vertebrates' heart there is, however, no well-defined bundle restricted to a small region, but a broad muscular connection or "funnel" bridging all or most of the auriculo-ventricular boundary. In the amphibia and reptiles the connecting fibers show a certain degree of differentiation, however, both macroscopically and microscopically (21). The fibers of the auriculo-ventricular funnel all pass in a down-

[16] Kent (105) has described an additional bundle of neuro-muscular structure passing across the auriculo-ventricular groove. This he later described as beginning above in a mass of "nodal tissue" lying in the right lateral auricular wall (106). After severing all auriculo-ventricular connections except this bundle, auricular contractions were found by Kent to be conducted to the ventricles (107). These findings, so far as we are aware, have not been confirmed, and opposed to them is the result of numerous investigators that severance of the His bundle alone results in auriculo-ventricular dissociation.

ward (caudal) direction. The funnel is interrupted by the valves posteriorly and the large vessels anteriorly, and the posterior (dorsal) part of the structure makes connection with the ventricular muscle higher than does the anterior portion. The elements of the funnel contain fewer fibrils and more sarcoplasm than the usual cardiac muscle. Further reduction in the funnel occurs until in mammals it is reduced to a small bundle lying anteriorly.

Gaskell (28) attempted to determine whether any special portion of the auriculo-ventricular ring in the tortoise heart was of greater importance for conduction than other parts, and found that section of the anterior part resulted in the greatest tendency toward dissociation. From similar experiments Laurens (109) concluded that the two lateral halves of the ring are more important, particularly the one on the left side, although in some cases the anterior portion alone can maintain conduction. Meek and Eyster (110) found that the anterior portion of the auriculo-ventricular ring was the first part to show the negative wave as the impulse passed to the ventricle.

The course of the wave of negativity in the ventricle of the tortoise during its activity was studied as long ago as 1883 by Burdon-Sanderson and Page (111), using the capillary electrometer. These investigators found that contacts on base and apex yielded diphasic curves indicating primary negativity of the base with passage to the apex and termination; in other words, a simple wave progressing from base to apex. This was confirmed by Waller (112), and Bayliss and Starling (113) came to the same conclusion as to the course of the wave in the mammalian heart. If the heart were cooled or was in otherwise bad condition, Bayliss and Starling found that a reversal of the normal direction of the electrical wave was the rule. The triphasic waves encountered in the mammalian heart were interpreted as indicating a maintenance of the condition of electrical negativity at the base after it had disappeared at the apex. The wave was supposed to begin at the base, as indicated by the upstroke on the curve, to pass to the apex and produce a downstroke on the curve, while the persistence of the condition of negativity of the base after its disappearance at the apex caused a final upstroke. Gotch (114) obtained similar triphasic curves from the tortoise and rabbit, but interpreted them as meaning a final return of the wave to the base. He believed that the negative wave appeared first at the auricular base, passed to the apex and returned to the aortic base, a course comparable to the direction in the primitive cardiac tube. Meek and Eyster (110) recorded the wave of negativity over the surface of the

tortoise ventricle by numerous contacts on the anterior, dorsal and lateral surfaces, and found the wave to spread around the auriculoventricular ring from the anterior part, thence to involve the whole base of the ventricle and finally to pass to the apex. No evidence of final base negativity was found. The third variation of the triphasic curve in base apex leads, comparable to the T wave of axial electrocardiograms, and interpreted by Gotch as indicative of terminal negativity at the base, was shown to be obtained from cardiac tissue in general without reference to the base of the ventricles. Clement (115) investigated the occurrence of the wave of negativity at points on the ventricular surface of the frog and tortoise, by comparing the incidence of the arrival of the wave at pairs of electrodes, the terminations of which were in contact.[17] He found the action current to appear on the surface almost simultaneously throughout, although the base and apex usually preceded the mid-region by a few thousandths of a second. Extrasystoles produced by artificial stimulation were conducted very much more slowly in the form of a wave spreading out from the point of stimulation. Erfmann (116), using the same method, compared the apex, middle, and right and left bases of the ventricles in dogs' and cats' hearts in situ. The base was found to receive the excitation a few thousandths of a second before the apex. The whole surface was regarded as receiving the impulse almost simultaneously. In the mammalian heart also an extra contraction produced by artificial stimulation was conducted with comparative slowness. Erfmann noted that the ventricular complex of the axial electrocardiogram began before (an average of about 0.035 second) the occurrence of the excitation wave on the ventricular surface, indicating that some internal portion received the excitation first. Using the same method on isolated artificially perfused dogs' hearts, Schneider (117) found the apex negative before the base, the reverse of that described by Erfmann in hearts in situ. The papillary muscles received the wave of negativity before the

[17] The two electrodes in each pair were connected by a thread of woolen yarn and the middle of this applied to the heart surface. A small part of the electrical change in the heart was conducted to the galvanometer, but sufficient to cause a deflection. By the use of electrodes of this type ("differential electrodes") it is thought that action currents arising from the activity of distant regions and conducted to the electrodes by the heart acting as an indifferent conductor are avoided, and that only changes are recorded due to tissue activity immediately beneath the electrodes. The difficulty with the use of this method arises from the necessity of working with a very loose galvanometer string in order to record the minute currents obtained.

apex (an averge of 0.012 second)[18] while the area showing initial negativity was found to be the interventricular septum, which preceded the earliest external surface contact by an average of 0.03 second.[19]

Lewis has reported a series of experiments on the hearts of toads, tortoises, birds and dogs, all in situ, in which the appearance and course of the wave of negativity was determined by the direction of movement of the galvanometer thread and by comparing the incidence in direct leads from various points with the R complex of the axial electrocardiogram.[20] Interior as well as surface points were compared. In the toad (125) the mid-region of the ventricle was the first part of the outer surface to show negativity, but this was preceded by the internal surface near the auriculo-ventricular ring. The latter coincided in time with the onset of the R complex, indicating further that it was the first part of the ventricle to receive the excitation wave. The apex preceded the base in three hearts, while the reverse was true in two. Surface readings differed by 0.015 to 0.035 second, indicating a much slower propagation than found by Clement. The truncus arteriosus was the last part to show indications of activity. It was concluded that the excitation spread from the internal surface of the auriculo-ventricular ring outward to the ventricular surface through the walls of the ventricle. In the tortoise (126) similar conditions were found, except that both

[18] Since the papillary muscles, in reference to the anatomical distribution of the auriculo-ventricular bundle, should be one of the earliest parts of the ventricles to receive the impulse, several investigators have attempted to record their contraction by mechanical suspension in isolated hearts under artificial perfusion. With the difficulties attending the recording of the contraction of a small mass of tissue attached to a larger moving mass and the doubtful interpretation of records thus obtained, it is not surprising that the conclusions should have been divergent. Roy and Adami (118) and Fenwick and Oberand (119) found the contraction of the papillary muscle follows contraction of the ventricles as a whole. The former worked with dogs, the latter with isolated rabbits' hearts. Hering (110) found the reverse to be true and Haycraft and Peterson (121) regarded the two as practically simultaneous. Finally, Salzman (122) found that the anterior papillary muscle contracted after the base but before the apex of the ventricle.

[19] De Meyer in 1907 (123) found the outer surface of the ventricle negative before the interior surface in excised frogs' and turtles' hearts, while Straub (124) found in artificially perfused tortoise hearts that the general interior surface (the perfusion solution forming a single electrode) was negative before the outside apical surface.

[20] Probable sources of error in such determinations have already been mentioned (page 23). Measurements with two pairs of electrodes in the heart surface are probably less subject to error.

the internal and external surfaces of the base were negative before the apex. In the bird Lewis (127) found that the whole process of spread through the ventricle occurred within 0.02 second, and from surface readings as well as electrocardiographic analysis (see page 33), concluded that the excitation spread downward through the interventricular septum to the apices and thence to the ventricular bases. More recent observations on the bird's heart by Mangold (38) in which surface comparisons of negativity were made, indicate that the excitation, at least on the outer surface, reaches the apex of the right ventricle and the base of the left ventricle first, and courses in opposite directions in the two ventricles. In the dog the first region on the outer surface of the ventricles to show evidence of reception of the excitation was found by Lewis and Rothschild (128) to be the ventral surface of the right ventricle near the mid-region of the septum. The wave passes from here to the right over the ventricle and to the base and apex. The left ventricle shows indication of surface activity first at the apex. Negativity appears almost simultaneously over a number of points quite widely separated, indicating that it is not conducted in the form of a wave through the superficial heart muscle but is distributed to the muscle through the conductive system (ramifications of auriculo-ventricular bundle and Purkinje network). Further support for this conclusion was obtained on section of the right branch of the auriculo-ventricular bundle, which resulted in no change in the time of arrival of the excitation over the left ventricle, while the contacts on the right ventricle showed a delay amounting to double or more the normal. Comparison of corresponding points on the interior or endocardial lining with the outer epicardial surface showed the former always to precede. The difference in time between these depended on the thickness of the muscular wall. The region of the ventricle showing first evidence of activity was found to be an area of the endocardium supplied by the initial branches of the left bundle (septum beneath aortic valve).[21] The main

[21] Lewis and Rothschild studied the time of arrival of the negative wave at a number of areas on the endocardial surface. With a few exceptions these readings showed differences of only a few thousandths of a second, probably within errors of measurement. These small differences preclude a satisfactory conclusion as to the direction of the spread of the wave on the endocardial surface, but do indicate that this spread is very rapid. The larger differences of reading on the epicardial surface are satisfactorily accounted for by the assumption, which these authors made, that the excitation spreads out from the Purkinje system through the usual ventricular musculature, and since the velocity is less in this, difference in thickness of the muscle mass would make considerable difference in the arrival of the wave at different points on the epicardial surface.

conclusion from this work was that the excitation wave is distributed throughout the endocardial surface by the branches of the bundle and the Purkinje network and passes from the terminations of this system to the musculature of the ventricles. The estimated velocity of the excitation wave in the special conductive system was from 3000 to 5000 mm. per second, a velocity approximately ten times as fast as that in the ordinary ventricular muscle.[22]

DEDUCTIONS AS TO THE COURSE OF THE EXCITATION WAVE IN THE VENTRICLE FROM ANALYSIS OF THE ELECTROCARDIOGRAM

Electrocardiograms, obtained by leading from electrodes applied to the surface of the body, are to be regarded as formed by a summation of the various action currents developed in the heart during the passage of the excitation from its point of origin to its final termination in the ventricles and the muscular contractions which follow in its wake. Einthoven in 1912 (130) called attention to the fact that if simultaneous points on the three standard electrocardiographic leads are known, and the direction and magnitude of the electrical change recorded by each are determined, it is possible to find the direction of the potential difference in the heart by using the method of the equilateral triangle.[23] This method was submitted to analysis by Einthoven, Fahr and De-Waart (131) and was applied to a determination of the direction of the electromotive force during different phases of the ventricular cycle by Fahr and Weber (132). The curves from the three leads give the

[22] Hering (75) attempted to measure the velocity of conduction in the auriculo-ventricular bundle and its branches by comparing the incidence between the stimulus and ventricular contraction when the ventricular end of the severed bundle was stimulated and when the ventricle was stimulated directly. No difference could be found. Erlanger (129) measured the velocity of conduction in the false tendons connecting the branches of the bundle and the papillary muscles, which histological examination showed are made up entirely of Purkinje tissue. The experiments were carried out on calves' hearts under artificial perfusion, and the rate of conduction was estimated by the difference in the incidence of ventricular response to two stimuli at different points on the tendon. The average figure obtained was 730 mm. per second, regarded by the author as probably lower than the true velocity, because measurement of the distance between the points stimulated was made on the flaccid tendon.

[23] The three standard electrocardiograph leads (lead I, two arms, lead II, right arm, left leg, lead III, left arm, left leg) form a sufficiently accurate equilateral triangle, since the potential difference between the two legs is so small that it can be neglected.

potential differences of the three corners of an equilateral triangle lying in the frontal plane of the body at every instant in the cardiac cycle, near the center of which lies the source of these potential differences, the heart. The direction of the potential differences within the triangle can be determined by plotting the value of the potential difference between corners (as determined by leads I and II, or II and III) at synchronous points along two sides, dropping perpendiculars from these and connecting the point of intersection with the vertex of the angle formed at the corresponding corner of the triangle. This line is then transferred to the center of the triangle. Its direction has the direction of the potential difference in the heart (electrical axis)[24] as manifested on the body surface at this moment, and its length expresses the so-called "manifest value"[25] of the potential. By similar plotting from instant to instant throughout the ventricular cycle, there is obtained a curve expressing the change in direction of the resultant potential difference and its magnitude throughout the progress of the cycle.[26] While such an analysis gives nothing more than the direction and magnitude of the resultant vector of potential differences in the heart, if considered in the light of our knowledge of the architecture of the conductive system, certain inferences may be drawn as to the path of the excitation in the ventricle.

Fahr and Weber (132) and Fahr (133) found that the initial resultant negativity, in most human hearts, corresponding to the Q wave in the electrocardiogram, is placed to the right of the center of the triangle and progresses in a clockwise direction through an angle of approximately 220° to the termination of the S wave of the electrocardiogram. In a smaller number, in which a Q wave occurs in leads I and II, initial resultant negativity points toward the left. There then follows a period in which the vectors are zero (space between S and T waves of electro-

[24] Currents flowing in other planes than the frontal are projected in this plane and their components corresponding to this plane form a part of the total value obtained. The three leads of electrocardiography lie in the frontal plane, so there is obtained only potential differences in this plane or projection upon this plane of potential differences lying in other planes of space.

[25] The manifest value is a magnitude formed by the geometrical summation in the frontal plane of the body of all the action currents present in the heart at that moment.

[26] Fahr and Weber found from experiments made on the cadaver, in which potential differences were induced in the heart by the direct application of a current from a storage battery, that the direction of potential difference, as indicated by the vector, was accurate within ± 10 degrees.

cardiogram), expressing the condition of negativity of the whole mass of muscle. The more important points in reference to the course of the excitation wave, as inferred from the above results, are as follows. Initial resultant negativity is in most cases on the right side, due probably to the short distance of the right bundle branch to the subendocardium of the right ventricle. Its first distribution in the right side is nearer apex than base. Within a short interval the negative wave has spread to the base of the right ventricle and to the left ventricle, and the vector then moves to the left of the midline, indicating preponderance of negativity in the left ventricular base. The spread to the left base is through two paths (two divisions of left bundle branch), which results in the excitation reaching a large mass of muscle at this time. Later the vector rotates toward the left apex with the involvement of this region, and the manifest value becomes very small, indicating that nearly the whole of the ventricular musculature has become involved.

Lewis (134) has analyzed the spread of the excitation separately in each ventricle of the dog's heart by the method of the equilateral triangle. In a previous paper Lewis and Rothschild (128) had shown that when one branch of the auriculo-ventricular bundle was cut, there was a delay of several hundredths of a second in the spread of the excitation to the corresponding ventricle, a time nearly or quite sufficient for the completion of the Q R S group of the electrocardiogram.[27] This delay is due to the fact that the normal path of the excitation to the ventricle is eliminated and the excitation spreads from that ventricle still retaining connection with the bundle across the interventricular groove. By cutting the branches to one or the other ventricle it was thus possible to obtain electrocardiograms signalling the course of the excitation

[27] Eppinger and Rothberger (135) and Rothberger and Winterberg (136) had previously sectioned the branches of the bundle in dogs and studied the changes in the axial electrocardiograms. They found that section of either branch resulted in a profound change in the curve, which assumed the unusual form characteristic of extrasystoles of the affected ventricle. They regarded the atypical electrocardiogram as due to the abnormal conduction to the ventricle separated from the bundle, the impulse spreading to it from the other ventricle and thus entering it and being conducted through it in an abnormal manner. Eppinger and Stoerk (137) published similar curves from five clinical cases supposedly due to right bundle branch defect, this conclusion being supported in two by examination at autopsy. Subsequent clinical examples were furnished by Lewis (138), by Mathewson (139), by Carter (140) and by Wilson and Hermann (141). Cohn and Lewis (142) reported that in four of the cases described by Carter, the pathological findings failed to support the diagnosis of bundle branch block.

through one ventricle alone, namely, the one retaining connection with the bundle, for a period comprising most or all that required for the spread of the excitation (period of Q R S group). Lewis designated the electrocardiograms obtained as levograms and dextrograms. Analysis of the dextrogram showed the vectors to take a clockwise path with initial resultant negativity at the base and terminal at the apex. The levogram showed on the other hand an anti-clockwise movement with initial resultant negativity near the mid-region and termination at the apex. These directions correspond with the inference of a spread through the initial branches of the bundle to the interventricular septum, the apices and around through the subendocardial Purkinje network on each side to the ventricular bases.[28] From each point in the Purkinje network during this spread the excitation is regarded as passing out through the ventricular muscle to reach the epicardial surface. By a combination of the levogram and dextrogram from the same heart, Lewis further found that a curve characteristic of the normal electrocardiogram was obtained, and that therefore if any two of these curves were at hand the third could be determined. He concluded that the electrocardiogram (Q R S group) is a "bigram" or summation curve of the action currents in the conduction system of the two ventricles (levogram and dextrogram).[29]

In a subsequent communication Lewis (145) analyzed the spread of the excitation by the method of the equilateral triangle in two human cases, one of supposed left, the other of supposed right bundle branch lesion. The former (dextrogram) showed a clockwise rotation of the vectors, in a general way similar to that obtained on section of the left bundle branch in dogs. The latter (levogram), showed an anti-clockwise movement, again resembling fairly closely the corresponding curve

[28] It should be pointed out that the method of the equilateral triangle has not been proven to be valid for the dog and the shape of the body makes it probable that it would not hold true.

[29] Selenin (143) had previously put forward a similar view from the results of the action of digitalis on the dog's electrocardiogram, and a study of the atypical curves obtained by extrasystoles occurring in the right and left ventricles respectively. He found that a combination of the atypical electrocardiogram from right and left ventricular extrasystoles gave an approximately normal electrocardiogram. According to Eiger (144), the electrocardiogram of each ventricle separately is a composite of two curves arising from the upward and downward passage of the excitation wave from the termination of the bundle branches, and the electrocardiogram of the whole heart is formed by a summation of these.

in dogs after section of the right bundle branch. Lewis also presented the data obtained from a large Rhesus monkey after experimental section of the right bundle branch with subsequent histological control. The vectors show an anti-clockwise rotation resembling the supposed levogram of the clinical cases.

DISCUSSION

Attempt has been made in this review to present the experimental facts as obtained by various investigators impartially so far as concerns any "theory" of the origin and conduction of the heart beat. A consideration of the experimental evidence we believe, however, lends strong support to the view that the property of impulse initiation and conduction is normally a function of the nodal tissues and the connecting anatomically differentiated structures. Initial activity, so far as we are able to determine it, is normally associated in both the lower vertebrate and mammalian heart with the sino-auricular node. Abolition of the function of this region results in the assumption of the "primum movens" by that part of the heart containing a similar mass of tissue (auriculo-ventricular node), separated a considerable distance from the normal seat of initial activity. Furthermore, most workers have found that the occurrence of all impulses arising within the heart under various abnormal or unusual circumstances (so-called ectopic beats), may be referred to one of the nodes or to some part of the auriculo-ventricular bundle. The only apparent exception to this is auricular fibrillation, in which the experimental evidence indicates that the impulses may be initiated in various portions of the auricular musculature. The type of contraction in this rhythm departs widely, however, from the contraction in other rhythms, in the absence of coördination between the individual muscle fibers. If this point of view is adopted, the myogenic-neurogenic controversy becomes a question to be decided by the histologist and the neurologist, and is a matter of the fundamental nature of the "neuro-muscular" tissues composing these structures. Recognition of this fact is perhaps no better expressed than by the great reduction of contributions to the literature, during the past ten years, concerning this prolonged and formerly active physiological controversy. According to this view, the tissues comprising the two nodes and the bundle with its ventricular connections, are to be regarded as that part of the heart which serves the function of initiating the excitation (property of automaticity or response to the "inner stimulus") and distributing it in a coördinated

manner to the musculature of the auricles and ventricles, the main function of which is contraction.[30]

The only line of modern experimental evidence uniformly opposed to the view of the sino-auricular node as the normal seat of impulse initiation, is that obtained from a study of the automaticity of isolated strips from the veno-auricular regions of the heart. These experiments, which have been quoted in preceding pages, have been carried out by Erlanger[31] and his co-workers. It would seem that a reasonable objection to the interpretation of these experiments in opposition to the view that nodal tissue is of paramount importance for automaticity, is to be found in the injury to which the heart tissue is subjected and the very abnormal conditions under which it is studied. The mechanical abuse associated with preparation of the strips, interruption of the vascular supply and exposure to artificial perfusion solutions, may very readily alter normal reactions and lead to the development of abnormal activity. There would seem to be no question that areas of the caval and right auricular regions outside of the sino-auricular node can develop automatic power as indicated by rhythmic beating. Skeletal muscle, too, may show evidences of rhythmicity under abnormal environmental conditions—(Loeb (1884), Dynamics of living matter, 1906, p. 82)—yet there is no evidence that this ever occurs in the living intact animal. As we have seen, most of the experimental evidence, it is true of a negative character by necessity of the case, indicates that only regions containing nodal tissue show automatic power in the intact heart under approximately normal conditions. We cannot but feel that the results from observations of cardiac strips might be held of considerable value if corroborating other experimental evidence, but are of little significance when such other evidence is uniformly in opposition.

[30] It is possible that differentiation of the properties of automaticity, conductivity and contractility in the mammalian heart has reached the point where certain strucures have entirely lost one property at the expense of the development of others to a pronounced degree. Eyster and Meek (33) found in dying hearts that the sulcus terminalis showed no visible contractions, while examination by the initial negativity method showed that the visible contractions in neighboring regions of the auricle and vena cava were resulting from excitations from this apparently quiescent region. Erlanger (129) was unable to find any evidence of contraction in false tendons composed of branches of the bundle, although the rate of conduction is faster in these than in ventricular muscle.

[31] An excellent presentation of the view which regards various portions of the veno-auricular region as normally possessing automatic power and against the recognition of the sino-auricular node as of paramount importance has been given by Erlanger (93).

Granted that the impulse arises normally within or in the immediate neighborhood of the sino-auricular node, the major point of issue is perhaps the question as to its manner of conduction to the auriculo-ventricular node. In regard to this it may be said that most anatomical work has failed to reveal a connection between the two nodes composed of tissue resembling the nodes or the auriculo-ventricular bundle, and it would seem probable that conduction between the two occurs by the usual cardiac tissue; however, the possibility of the satisfactory confirmation of those workers who have found an anatomically differentiated connection is not to be excluded. From the experimental standpoint, all observations seem to agree that conduction is rather diffuse, or at least not limited to a narrow path such as exists between the auriculo-ventricular node and the ventricles. According to one view, this path is very diffuse, and occurs through the right atrium, while according to another view the excitation passes to the atrium and auriculo-ventricular nodes separately, although passage through the atrium may occur when the spread in the normal direction is prevented. This latter view assumes two or more possible paths of unequal resistance, the excitation normally taking the path of least resistance.

In the frog and tortoise the excitation spreads to the ventricle by way of the fibers passing downward from the auriculo-ventricular ring, the so-called funnel. The conduction path is broad and the excitation presumably passes down all the fibers of the funnel, although there is evidence that not all the funnel is necessary for conduction and perhaps certain parts are more important in this connection than others. The excitation passes to the ventricular muscle and most workers agree that it proceeds as a simple wave to the apex, where it terminates.

Delay in the spread of the excitation from the supra-ventricular regions to the ventricles in the mammalian heart apparently occurs, in larger part, in the auriculo-ventricular node. Having passed the node, the excitation is distributed through the auriculo-ventricular bundle, and probably reaches the ventricular musculature through the extensive subendocardial Purkinje network. The spread through this system is rapid, so rapid indeed that it is difficult or impossible to determine very definitely the direction taken by the excitation. Due to the extensive branching of the auriculo-ventricular conducting system the excitation probably spreads in various directions, and no longer proceeds as a single wave along a circumscribed path. Determinations of the incidence of negativity on the endocardial surfaces have given differences of only a few thousandths of a second. According to some

investigators this is also true of the epicardial surface. Others find, however, that while many epicardial surface points receive the excitation at approximately the same time, there are other regions which do not become excited until a few hundredths of a second later. This slower conduction to the latter regions may be due to a longer conducting path or to a slow rate of propagation through the ordinary musculature after the excitation leaves the Purkinje system, thus reaching the outer surface over regions where the ventricular muscle is thin before it reaches the surface areas overlying thicker masses of muscle. All recent workers seem to agree that in the mammalian ventricle the excitation cannot be regarded as a simple wave propagated in a single path, but it is to be regarded as coursing in various directions through the complex ramifications of the conduction system.

BIBLIOGRAPHY

(1) His, W.: Abhandl. d. Sachs. Ges. d. Wiss. Math.-Physik.-Kl, 1893, xix, 1.
(2) His, W.: Zentral. f. Physiol. 1895, ix, 469; Deutsch. Arch. f. klin. Med., 1899, lxiv, 316.
(3) Humblet, M.: Arch. internat. d. Physiol., 1904, i, 278.
(4) Hering, H. E.: Arch. f. d. gesammt. Physiol., 1905, cviii, 97.
(5) Erlanger, J.: Zentralbl. f. Physiol., 1905, xix, 9; Journ. Exper. Med., 1906, viii, 51.
(6) Tawara, S.: Das Reizleitungssystem des Saügetierherzen, Jena, 1906.
(7) Keith, A., and Flack, M.: Journ. Anat. and Physiol., 1907, xli, 172.
(8) Wenckebach, K. F.: Arch. f. (Anat. u.) Physiol., 1906, 297.
(9) Morrison, A.: Journ. Anat. and Physiol., 1912, xlvi, 319.
(10) Oppenheimer, B. S., and Oppenheimer, A.: Journ. Exper. Med., 1912, xvi, 613.
(11) Nagayo: Verhandl. d. deutsch. path. Gesellsch., 1908, xii, 150.
(12) Koch, W.: Deutsch. med. Wochenschr., 1910, xxxvi, 688.
(13) Koch, W.: Deutsch. med. Wochenschr., 1909, xxxv, 429.
(14) Koch, W.: Zentralbl. f. allg. Path. u. path. Anat., 1909, xx, 487.
(15) Koch, W.: Med. Klinik, 1912, viii, 108.
(16) Koch, W.: Arch. f. d. gesammt. Physiol., 1913, cli, 279.
(17) Keith, A.: Lancet, 1904 (I) 555, 629 and 703.
(18) Keith, A., and Mackenzie, I.: Lancet, 1910, lxxxviii, 101.
(19) Mackenzie, I.: Zentralbl. f. allg. Path. u. path. Anat., 1910, xxi, 433.
(20) Mackenzie, I.: Journ. Path. and Bact., 1910, xiv, 404.
(21) Külbs, F.: Das Reizleitungssystem im Herzen, Berlin, 1913.
(22) Aschoff, L.: Zentralbl. f. allg. Path. u. path. Anat., 1910, xxi, 433.
(23) Aschoff, L.: Deutsch. med. Wochsenchr., 1914, xl, 1036.
(24) Mall, F. P.: Amer. Journ. Anat., 1912, xiii, 249.
(25) Curran, E. J.: Anat. Anzeiger, 1909 xxxv, 89.
(26) Thorel: Zentralbl. f. allg. Path. u. path. Anat., 1911, xxi, 433.

(27) Aschoff, L., and Moncheberg, K.: Zentralbl. f. allg. Path. u. path. Anat., xxi, 433.
(28) Gaskell, W. H.: Journ. Physiol., 1883, iv, 43.
(29) Englemann, T. W.: Arch. f. d. gesammt. Physiol., 1896, lxv, 115.
(30) MacWilliam, J. A.: Journ. Physiol., 1885, vi, 119.
(31) Garrey, W. E.: Amer. Journ. Physiol., 1911, xxviii, 330.
(32) Meek, W. J., and Eyster, J. A. E.: Amer. Journ. Physiol., 1912, xxxi, 31.
(33) Meek, W. J., and Eyster, J. A. E.: Amer. Journ. Physiol., 1916, xxxix, 291.
(34) Schlomovitz, B. H., and Chase, C. S.: Amer. Journ. Physiol., 1916, xli, 112.
(35) Flack, M.: Arch. internat. d. Physiol., 1911, xi, 121.
(36) Mangold, E., and Kato, T.: Arch. f. d. gesammt. Physiol., 1914, clvii, 1.
(37) Mangold, E.: Deutsch. med. Wochenschr., 1914, xl, 1035; Die Erregungsleiten im Wirbeltierherzen, Jena, 1914.
(38) Mangold, E.: Arch. f. d. gesammt. Physiol, 1919, clxxv, 327.
(39) Zahn, A.: Arch. f. d. gesammt. Physiol, 1913, cxli, 247.
(40) MacWilliam, J. A.: Journ. Physiol., 1888, ix, 167.
(41) Gaskell, W. H.: Philosoph. Trans. Roy. Soc. London, 1882, iii, 993.
(42) Adam: Arch. f. d. gesammt. Physiol., 1906, cli, 607.
(43) Langendorff, O., and Lehmann: Arch. f. d. gesammt. Physiol., 1906, cxii, 352.
(44) Langendorff, O.: Arch. f. d. gesammt. Physiol., 1906, cxii, 522.
(45) Erlanger, J., and Blackmann, J. R.: Amer. Journ. Physiol., 1907, xix, 125.
(46) Erlanger, J.: Amer. Journ. Physiol., 1910, xxvii, 87.
(47) Hering, H. E.: Arch. f. d. gesammt. Physiol., 1907, cxvi, 143.
(48) Hering, H. E.: Arch. f. d. gesammt. Physiol., 1900, lxxxii, 1.
(49) Fredericq, L.: Arch. internat. de Physiol., 1906, iv, 160.
(50) Hering, H. E.: Münch. med. Wochenschr. 1909, lvi, 845.
(51) Hirschfelder, A. D., and Eyster, J. A. E.: Amer. Journ. Physiol., 1907, xviii, 221.
(52) Koch, W.: Beitr. z. path. Anat. u. z. allg. Path., 1907, xlii, 203.
(53) Eyster, J. A. E., and Meek, W. J.: Heart, 1914, v, 137.
(54) Fredericq, L.: Acad. r. d. Belg. Bull., Cl. des Sciences, 1901, 125.
(55) Hering, H. E.: Arch. f. d. gesammt. Physiol., 1905, cvii, 97.
(56) Fredericq, L.: Arch. internat. d. Physiol., 1912, xii, 109.
(57) Rehfish, E.: Deutsch. med. Wochenschr., 1909, xxxvi, 977.
(58) Eyster, J. A. E. Journ. Exper. Med., 1910, xii, 257.
(59) Eyster, J. A. E., and Meek, W. J.: Arch. Int. Med., 1913, xi, 204.
(60) Englemann, T. W.: Arch. f. (Anat. u.) Physiol., 1903, 505.
(61) Lohmann, A.: Arch. f. (Anat. u.) Physiol., 1904, 431.
(62) Mackenzie, J.: Brit. Med. Journ., 1904, (i), 529.
(63) Hering, H. E. Zentralbl. f. Physiol., 1905, xix, 129.
(64) Lohmann, A.: Arch. f. d. gesammt. Physiol., 1908, cxxiii, 628.
(65) Burdon-Sanderson, J., and Page, F. J. M.: Journ. Physiol., 1879, ii, 384; 1883, iv, 327.
(66) Waller, A.: Journ. Physiol., 1887, viii, 229.
(67) Bayliss, W. M., and Starling, E. H.: Journ. Anat. and Physiol., 1892, ix. 207.

(68) Tawara, S.: Zentralbl. f. Physiol., 1905, xix, 70.
(69) Jaeger, T.: Deutsch. Arch. f. klin. Med., 1910, c, 1.
(70) Flack, M.: Journ. Physiol., 1910, xli, 64.
(71) Flack, M.: Arch. internat. d. Physiol., 1911, xi, 111, 119 and 127.
(72) Mackenzie, I.: Zentralbl. f. allg. Path. u. path. Anat., 1910, xxi, 433.
(73) Hering, H. E.: Arch. f. die gesammt. Physiol., 1910, cxxxvi, 466.
(74) Hering, H. E.: Zeitschr. f. exper. Path. u. Therap., 1906, ii, 521.
(75) Hering, H. E.: Arch. f. d. gesammt. Physiol., 1910, cxxxi, 572.
(76) Magnus-Alsleben: Arch. f. exper. Path. u. Pharm., 1911, lxiv, 228.
(77) Cohn, A. E., and Kessel, L.: Arch. Int. Med., 1911, vii, 226.
(78) Cohn, A. E., Kessel, L., and Mason, H. H.: Heart, 1912, ii, 311.
(79) Moorehouse, V. H. K.: Amer. Journ. Physiol., 1912, xxx, 358.
(80) Brandenburg, K., and Hoffmann, P.: Med. Klinik, 1912, viii, 16.
(81) Ganter, G., and Zahn, A.: Zentralbl. f. Physiol., 1912, xxv, 782; Arch. f. d. gesammt. Physiol., 1912, cxlv, 335.
(82) Schlomovitz, B. H., and Chase, C. S.: Arch. Int. Med., 1917, xx, 613.
(83) Englemann, T. W.: Arch. f. d. gesammt. Physiol., 1897, xv, 109.
(84) Sansum, W. D.: Amer. Journ. Physiol., 1912, xxx, 421.
(85) Moorehouse, V. H. K.: Amer. Journ. Physiol., 1913, xxi, 421.
(86) Wybauw, R.: Bull. d. l. Soc. Roy. d. Sc. d. Brux., 1910, no. 5; Arch. internat. d. Physiol., 1910, x, 78.
(87) Lewis, T., Oppenheimer, B. S., and Oppenheimer, A.: Heart, 1910, ii, 147.
(88) Lewis, T.: Heart, 1909, i, 306.
(89) Kölliker and Müller: Wurz. Verhandl., 1856, vi, 528.
(90) Lewis, T.: Heart, 1920, vii, 131.
(91) Lewis, T.: Heart, 1910, ii, 23.
(92) Eyster, J. A. E., and Meek, W. J.: Heart, 1914, v, 119.
(93) Erlanger, J.: Arch. Int. Med., 1913, xi, 204.
(94) Garten, S.: Skand. Arch. f. Physiol., 1913, xxix, 114.
(95) Lewis, T., Meakins, J., and White, P. D.: Phil. Trans. Roy. Soc. London, 1914, Ser. B, ccv, 375.
(96) Meek, W. J., and Eyster, J. A. E.: Heart, 1914, v, 227.
(97) Meek, W. J., and Eyster, J. A. E.: Amer. Journ. Physiol., 1914, xxxiv, 368.
(98) Eyster, J. A. E., and Meek, W. J.: Arch. Int. Med., 1916, xviii, 775.
(99) Eyster, J. A. E., and Meek, W. J.: Arch. Int. Med., 1917, xix, 117.
(100) Eyster, J. A. E., and Meek, W. J.: Heart, 1912, iv, 137.
(101) Eyster, J. A. E., and Evans, J. S.: Arch. Int. Med., 1915, xvi, 832.
(102) Levine, S. A.: Arch. Int. Med., 1916, xvii, 153.
(103) Naish, A. E.: Quart. Journ. Med., 1914. viii, 33.
(104) Eyster, J. A. E., and Meek, W. J.: Amer. Journ. Physiol., 1917, xlii, 611.
(105) Kent, S. A. F.: Journ. Physiol., 1893, xiv, xxiv.
(106) Kent, S. A. F.: Brit. Med. Journ., 1914 (ii), 105.
(107) Kent, S. A. F.: Journ. Physiol., 1914, xlviii, cvii.
(108) Lhamon, R. M.: Amer. Journ. Anat., 1912, xiii, 55.
(109) Laurens, H.: Arch. f. d. gesammt. Physiol., 1913, cl, 139.
(110) Meek, W. J., and Eyster, J. A. E.: Amer. Journ. Physiol., 1912, xxxi, 31.

(111) Burdon-Sanderson, J., and Page, F. J. M.: Journ. Physiol., 1883, iv, 327.
(112) Waller, A.: Phil. Trans. Roy. Soc. London, 1889, Ser. B., clxxx, 169.
(113) Bayliss, W. M., and Starling, E. H.: Internat. Monatschr. f. Anat. u. Physiol., 1892, ix, 256.
(114) Gotch, F.: Heart, 1909, i, 235.
(115) Clement, S.: Zeitschr. f. Biol., 1912, lviii, 110.
(116) Erfmann, W.: Zeitschr. f. Biol., 1913, lxi, 155.
(117) Schneiders, M.: Zeitschr. f. Biol., 1915, lxv, 465.
(118) Roy and Adami, J. G.: The practitioner, 1890, xliv, 88.
(119) Fenwick, W. S., and Overend, W.: Brit. Med. Journ., 1891, (i), 1117.
(120) Hering, H. E.: Deutsch. med. Wochenschr., 1907, xxxiii, 1895.
(121) Haycraft, J. B., and Peterson, D. R.: Journ. Physiol., 1896, xix, 262.
(122) Saltzman, F.: Skand. Arch. f Physiol., 1908, xx, 233.
(123) DeMeyer, J.: Arch. internat. de Physiol., 1907, v, 76.
(124) Straub, H.: Zeitschr. f. Biol., 1909, liii, 499.
(125) Lewis, T.: Phil. Trans. Roy. Soc. London, 1915, Ser. B, ccvii, 221.
(126) Lewis, T.: Phil. Trans. Roy. Soc. London, 1915, Ser. B, ccvii, 240.
(127) Lewis, T.: Phil. Trans. Roy. Soc. London, 1915, Ser. B, ccvii, 298.
(128) Lewis, T., and Rothschild, M. A.: Phil. Trans. Roy. Soc. London, 1915, Ser. B, ccvi, 181.
(129) Erlanger, J.: Amer. Journ. Physiol., 1912, xxx, 395.
(130) Einthoven, W.: Lancet, 1912, (i) 853.
(131) Einthoven, W., Fahr, G., and DeWaart, A.: Arch. f. d. gesammt. Physiol., 1913, cl, 275.
(132) Fahr, G., and Weber, A.: Deutsch. Arch. f. klin. Med., 1915, cxviii, 361.
(133) Fahr, G.: Arch. Int. Med., 1920, xxv, 146.
(134) Lewis, T.: Phil. Trans. Roy. Soc. London, 1915, Ser. B, ccvii, 247.
(135) Eppinger, H., and Rothberger, C. J.: Zeitschr. f. klin. Med., 1910, lxx, 1.
(136) Rothberger, C. J., and Winterberg, H.: Zentralbl. f. Physiol., 1911, xxiv, 1055; Zentralbl. f. Herz. u. Gefasskrankh., 1913, v, 206.
(137) Eppinger, H., and Stoerk, O.: Zeitschr. f. klin. Med., 1910, lxxi, 157.
(138) Lewis, T.: Clinical electrocardiography, London, 1913.
(139) Mathewson, G. D.: Heart, 1913, iv, 385.
(140) Carter, E. P.: Arch. Int. Med., 1914, xiii, 803.
(141) Wilson, F. N., and Herman, C. R.: Arch. Int. Med., 1920, xxvi, 153.
(142) Cohn, A. E., and Lewis, T.: Proc. N. Y. Path. Soc., 1914, xiv, 207.
(143) Selenin, W. P.: Arch. f. d. gesammt. Physiol., 1912, cxliii, 137.
(144) Eiger, M.: Arch. f. d. gesammt. Physiol., 1913, cli, 37; 1915, clxii, 433.
(145) Lewis, T.: Phil. Trans. Roy. Soc., London, 1915, Ser. B, ccvii, 284.

THE PRESENT STATUS OF THE PROBLEMS OF ANAPHYLAXIS

H. GIDEON WELLS

From the Otho S. A. Sprague Memorial Institute and the Department of Pathology of the University of Chicago

The literature of anaphylaxis and related topics, and the number of unsettled problems, are both too vast to permit of an exhaustive review within the scope contemplated by these PHYSIOLOGICAL REVIEWS, even if there were not already numerous excellent and adequate compilations of this sort in available literature (1). In order to avoid covering again too much of the ground already thoroughly surveyed by others, this review will limit itself to discussing what seem to be the most essential problems in the light of the more recent contributions.

DEFINITION AND LIMITATION OF THE TERM "ANAPHYLAXIS"

Much unnecessary work and many misleading statements would have been avoided in the past if more attention had been given to these essential points. Not all sudden attacks of dyspnea and convulsions in injected guinea pigs, or of cardiac syncope or cutaneous irritation in man, are manifestations of anaphylaxis, or are even related to this condition beyond in a greater or less similarity of these particular symptoms. Even distention of the lungs in guinea pigs, circulatory disturbances in rabbits, or splanchnic congestion in dogs, following intravascular injections of foreign materials, are not necessarily proof of a true anaphylactic reaction, for each and all may be produced by materials or under conditions that surely do not involve the same processes that are responsible for the reaction which occurs when a small quantity of an otherwise harmless protein is injected into an animal that has been previously prepared at a sufficiently great interval of time, by injection of the same protein. We must grant that this last is true anaphylaxis, and must demand that the term anaphylaxis be limited to conditions that result from a similar mechanism within the body of the reacting animal. At least the following criteria must be met:

1. The observed toxicity of the injected material must depend upon the sensitization of the animal; i.e., the substance must not produce similar symptoms in non-sensitized animals.

2. The symptoms produced must be those characteristic of anaphylactic intoxication as observed in the usual reactions with typical soluble proteins, being therefore the same for all antigens with the same test animal, but differing characteristically with each species of animal.

3. It should be possible to demonstrate passive sensitization with the serum of sensitized animals.

4. It should be possible to demonstrate typical reactions in the virgin guinea pig uterus strip.

5. It should be possible to demonstrate amelioration or prevention of the bronchial spasm in guinea pigs by proper use of atropin and epinephrin.

6. The possibility that the observed symptoms are caused by capillary thrombosis or embolism must be excluded.

7. After recovery from anaphylactic shock there should be exhibited a condition of desensitization under proper conditions.

We know that the foregoing conditions are fulfilled in the typical anaphylactic reaction with soluble protein antigens. Reactions produced by substances not coming within this class cannot properly be considered as true anaphylactic reactions unless they can be demonstrated to exhibit the same phenomena. The term allergy is broad enough to cover all those manifestations of altered reactivity and to indicate their relationship to one particular form of allergy, anaphylaxis.

An important series of papers by Karsner and Hanzlik (2) demonstrates the common sources of error in much of the published work on anaphylaxis, and especially on the supposed formation of anaphylatoxin in vitro by other than antigen-antibody reactions. It will be recalled that Friedberger, Bordet, Novy and DeKruif, and others, have observed symptoms considered to be characteristic of anaphylaxis following intravenous injection into animals of serum that has been acted upon by colloids, such as agar and inulin, or suspensions, such as kaolin or starch, or after injection of these same colloids or suspensions. Likewise, anaphylactic shock is said to have followed intravenous injections of acacia, arsphenamine and many other substances used in human therapeutics. In most of the animal experimentation the proof of the anaphylactic character of the reaction has rested not only on symptomatology but also on the presence of apparently typical bronchial spasm with persistent pulmonary distention in guinea pigs, or fall in blood pressure with splanchnic congestion in dogs.

Karsner and Hanzlik show that agar and numerous other substances produce "anaphylactoid" symptoms after intravenous injection, by virtue of their causing occlusion of the pulmonary capillaries by thrombosis or by agglutinated corpuscles or platelets. With such capillary obstruction, which is demonstrated both by finding the emboli and by the stoppage of perfusion fluids injected into the pulmonary vessels, there may occur a marked bronchial constriction which causes asphyxia and permanent distention of the lungs. This bronchial constriction, which has commonly been interpreted as proof of a true anaphylactic shock, differs from that of true anaphylaxis in not depending on a contraction of the bronchial musculature, which they contend is shown by the failure of epinephrin and atropin to influence the respiratory effects of agar in the same way that they control anaphylactic bronchial spasm. "On the basis of the results obtained with atropin and epinephrin, the mechanism of the action of agar and similar agents bears no relationship to true anaphylaxis or anaphylactic shock." How the "passive broncho-constriction" of agar intoxication is produced is not altogether satisfactorily explained.

> The agar forms thrombi which block the small pulmonary vessels, surrounding the bronchioles, which in turn are compressed by the *vis a tergo* of the pulmonary flow from the right heart of the intact animal, and of the perfusion fluid in surviving lungs. In the perfused lung the flow of perfusion fluid stops. The compressed bronchioles are sufficiently narrowed in caliber to prevent the escape of the air forced in by the violent efforts of inspiratory dyspnoea of the intact animal and the artificial respiration of the perfused lungs. In other words there is a passive broncho-constriction (really a compression of the bronchi), and as a result permanent inflation of the lung. . . . . It is conceivable that additional broncho-constriction is superimposed as a result of central asphyxial stimulation of the vagus (parasympathetic) on the passive constriction primarily brought about.

The capacity of agar to produce agglutination of red corpuscles is very marked. Concentrations as low as 0.001 per cent cause agglutination in vitro, while the corresponding figures for other colloids (e.g., acacia, 1.0; althea, 0.08; collargol, 0.005; gelatin, 0.05; nuclein, 0.5; beef serum, 0.005 per cent) indicate their capacity to cause anaphylactoid symptoms when injected intravenously. Therefore it is not surprising that on intravenous injection "the quantity of agar necessary to produce profound effects and even death, is remarkably small, namely, about 0.014 to 0.05 mgm. per gram of body weight, or about 0.014 to 0.025 per cent concentration of agar in the blood."

Perhaps it should be added that "no emboli, thrombi or conglutinated corpuscles were demonstrable in perfused or intact lungs of anaphylactic shock, after peptone, certain serums and histamine. All of these agents were also active smooth muscle stimulants in other regions besides bronchi, namely uterus and intestine."

With the above facts established it must be recognized that much of the work published and interpreted as anaphylaxis deals rather with anaphylactoid phenomena. In working with vegetable proteins, the danger of confusion of capillary embolism is particularly marked, since many plant products are powerfully hemagglutinative, especially extracts from the leguminous seeds, while many vegetable proteins are thrown out of solution by serum.[1] I have, therefore, usually selected the intraperitoneal route with the view of avoiding these sources of error. Karsner and Hanzlik have not investigated the amount of obstruction to pulmonary capillaries that may follow intraperitoneal injection of foreign colloids. Presumably it is much less, and perhaps entirely absent, but they do mention some slight pulmonary symptoms following subcutaneous injections of thromboplastic agents. Certainly in anaphylactic investigations, especially when the intravenous injection is used, the burden of proof of the nature of observed reactions lies heavily upon the investigator.

NATURE OF THE ANTIGENS (ANAPHYLACTOGENS)

It now seems to be generally accepted that true anaphylaxis represents an antigen-antibody reaction, but as yet there is not such complete agreement as to what the antibody is or as to what substances may behave as antigens in this particular reaction, i.e., anaphylactogens. There is no doubt that any soluble protein which can act as an antigen in other immunological reactions can act as an anaphylactogen, and the only soluble proteins as yet found not to be antigenic are those which are, from the chemical sense, incomplete proteins. As might be expected, such proteins as have a very small variety of amino acids are the chief examples of the non-antigenic proteins, notably the protamines and histones (3); also the commonly-occurring complexes of these with non-protein radicals, notably, the nucleoproteins (4) or hemoglobins (5). However, so complex and soluble a

[1] Undoubtedly the results obtained with the juice of *Cotyledon scheideckeri* by Kritchewsky (94) are merely anaphylactoid phenomena from a hemagglutinative and precipitating plant extract.

protein as gelatin seems to have absolutely no antigenic power, whether tested by anaphylaxis or by more sensitive methods, as shown conclusively by the careful work of Strain (6). As the chief known difference between gelatin and the proteins that do exhibit antigenic properties lies in an almost complete absence of tryptophane and tyrosine and a very low phenylalanine content, it would seem probable that it is this deficiency in aromatic amino acids that is responsible for the lack of antigenic capacity. Furthermore, there do exist proteins which lack one or more amino acids which are commonly present in typical proteins, yet which are strongly antigenic (e.g., zein from corn, which lacks glycine and tryptophane; gliadin of wheat and rye, or hordein of barley, which contain no glycine or lysine and very little arginine or histidine) (7). The fact that these three last proteins, which are so extremely poor in diamino acids, are potent, while the protamines which consist chiefly of diamino acids, are inert, suggests that these three diamino acids are not of importance in the anaphylactic activity of proteins, but since no protein is known which does not contain either histidine or arginine, we cannot prove this point as it seems to be proved for lysine.

Solubility of the protein is essential, since although insoluble proteins may eventually be brought into solution in the animal body, the process is too slow to bring about reactions, and also it is probably accompanied by disintegration of the protein molecule. Heat, of degrees that do not disintegrate the proteins, affects them only to the extent that it makes them insoluble. There are, however, only a few known proteins that are not made insoluble by heat, and these, except gelatin, are antigenic despite boiling, these being casein (8), ovomucoid (9), the so-called proteoses of plant seeds (10), beta-nucleoproteins (4) and perhaps the capsular substance of pneumococci (11). Compound proteins will produce anaphylaxis if they are soluble and if the protein constituents are themselves antigenic, but apparently not when the protein radical is a non-antigenic histon or protamin. Thus, alpha-nucleoproteins (4) and hemoglobin (5) are non-antigenic, while mucins (12), beta-nucleoproteins (4) and hemocyanin (13) are antigenic.

A suggestive observation on the requirements necessary for anaphylactogenic activity is furnished by Ten Broeck's observations (14) with the racemized proteins of Dakin. These proteins, presumably because of changes in structural configuration, are not attacked by the ordinary proteolytic enzymes, and likewise they do not exhibit anaphylactogenic properties. Landsteiner's (15) studies on the precipitin

reaction with azotized proteins have led him to the conclusion that the significance of the proteins for the usual immunity reactions, except for their immunizing capacity, does not depend so much upon their particular chemical structure as on their molecular size and the properties depending thereon. Such large molecular complexes may be able to react with antibodies even when incapable of engendering their formation.

The question whether anything less than an entire protein molecule can act as an anaphylactogen has been the subject of much investigation. Although I have always obtained negative results with numerous products of protein cleavage, nevertheless Fink (16), making in my laboratory an extensive study of proteose fractions obtained from coagulated egg white by hydrolysis with steam under pressure, did find evidence of a very slight anaphylactogenic property in those fragments that were precipitated at $\frac{3}{4}$ and complete saturation with ammonium sulphate, but not with fractions precipitated at $\frac{1}{4}$, $\frac{1}{3}$, $\frac{1}{2}$ and $\frac{2}{3}$ saturation. However, these reactions consisted only of slight symptoms even after intracardiac injections of 25 to 50 mgm. of proteose, and were of significance chiefly because similar symptoms were not produced by the other fractions and because positive precipitin and complement fixation reactions were obtained with these same two fractions only. This result is particularly surprising since it would be expected that the fractions precipitated by smaller concentrations of ammonium sulphate would represent larger molecules more closely resembling the intact protein molecule.

Fink gives an elaborate critical review of the published work on antigenic properties of the cleavage products of proteins, which indicates mainly that much of it is worthless, and that there is no conclusive evidence of anaphylactic or other immunological reactions having been produced with proteoses and peptones, beyond the very slight positive evidence that he has secured with his own carefully prepared materials. The obvious source of error in much of this work lies in the fact that "proteose" and "peptone" preparations generally are themselves active stimulators of smooth muscle and perhaps capillary poisons (Dale), in animals that are not sensitized, and hence observed symptoms can differ in sensitized and nonsensitized animals only quantitatively. They readily produce anaphylactoid manifestations on intravenous injection. Corroborating this interpretation of the supposed positive anaphylaxis reactions with proteoses is the fact that nearly all attempts to produce precipitins or complement-binding anti-

bodies with the same protein cleavage products have given negative results.

It is appalling to observe the amount of uncritical work done with unpurified, unanalyzed commercial "peptones," usually Witte's products. Such work can produce absolutely no usable evidence concerning most of the problems under investigation, since the source and composition of the material is entirely unknown to the investigator, and it has been repeatedly shown that Witte's peptone is a mixture of widely varying composition in different samples.

If no positive results can be obtained with most of the fractions of protein cleavage, including digestion mixtures of varying sources, but containing usually all the cleavage products from proteoses through peptones and polypeptids to amino acids, it is not easy to accept the statement that anaphylaxis can be produced by synthetic polypeptids, even one containing 14 molecules of leucine and glycine (Abderhalden, 17), still less the positive results of Zunz (18) with much simpler polypeptids (3–5 glycylglycine). These reactions were obtained after intravenous injections of the peptids into sensitized rabbits, and consisted of fall of blood pressure, increased rate of respiration and expulsion of feces and urine, but these reactions, slight as they are, were by no means constant, and occurred under very particular conditions, e.g., positive results with animals sensitized with six intraperitoneal injections at weekly intervals, but not in animals that received six subcutaneous injections at similar intervals. No other species of animals seems to have been tried, and it seems most improbable that the reactions observed would fulfil the requirements for proof laid down in the opening section of this review. Certainly, in view of the possible sources of such reactions independent of true anaphylactic shock, and the large number of negative results with peptones and digestion mixtures, the acceptance of the results yet obtained with synthetic polypeptids as proof of their antigenic activity will require much more convincing evidence than any that has yet been produced.

There still remains no satisfactory proof that anything except proteins can act as anaphylactogens. Even some definite antigens, whose protein nature is doubtful, are not anaphylactogens in the strict sense, namely, true soluble toxins and tuberculin, despite the fact that they do produce a certain type of hypersensitiveness (see Coca, 1). While many attempts have been made to establish the antigenic activity of lipoids by means of the precipitin and complement fixation reactions, the lack of solubility of the lipoids has made them unsuitable for ana-

phylaxis studies. Although a few studies with such material have been reported to give positive results, it is probable that these are all erroneous (see White, 19). Certainly as yet the antigenic capacity of lipoids has not been established.

On the other hand drugs and chemicals of many sorts have produced both local and systemic reactions that have been interpreted as anaphylaxis. Especially is this true of mercurials, salvarsan and other arsenicals, iodin and bromine and compounds thereof; also innumerable alkaloids (e.g., quinine, atropin) synthetic organic compounds (salicylates, antipyrine) and various vegetable oils, resins, etc. The evidence concerning the reactions produced by these substances has been thoroughly reviewed by Coca, who concludes that their anaphylactic nature has not been established. Wolff-Eisner has advanced the hypothesis that drug reactions may be produced by the blood or tissue proteins altered by the chemical so that they are as foreign proteins to the injected animal. As yet this hypothesis is unproved, although it is worth more extended and careful investigation than it has yet received. It is probable that many of the systemic reactions observed after intravenous injections of arsenicals and other substances are not anaphylactic but anaphylactoid, and dependent upon capillary thrombosis, agglutination emboli, or capillary toxicity, after the order of the results obtained by Karsner and Hanzlik. The failure to produce passive anaphylaxis with the serum of persons hypersensitive to drugs, the failure to secure convincing active sensitization in guinea pigs, the number of instances in which drug hypersensitiveness is exhibited on the first injection and those in which heredity is evidently responsible, and the numerous points of difference from true anaphylactic reactions, all make the identity of drug hypersensitivity with true antigen-antibody anaphylaxis, a matter of great improbabiliy.

THE NATURE OF THE IMMUNE-BODY

Anaphylactin, sensibilizin, sensitizin

The demonstration that anaphylactic sensitization can be transferred to a normal animal by injection of blood or serum from a sensitized animal, has been generally accepted as establishing anaphylaxis as a typical antigen-antibody reaction. Doerr and Russ added important support to this view by showing that the potency of a serum in conferring passive sensitization varies in direct proportion to its content in precipitins. Zinsser states that "The principle that anaphylaxis

depended directly on the meeting of the antigen with its specific antibody has never been seriously questioned since that time." Three main questions have remained unsolved, however, namely: *a*, Is the sensitizing antibody the same as the precipitin or the complement-fixing antibody? *b*, Where does the reaction of antigen and antibody take place? *c*, How does this reaction produce the manifestations of anaphylactic shock?

THE RELATION OF PRECIPITIN TO THE SENSITIZING ANTIBODY

An attractive hypothesis is that there are two fundamental types of immunity reactions. One, having to do with substances which are essentially active poisons, neutralizes or inhibits their toxicity by direct chemical action (see Zinsser, 1). In this group come the antitoxins and the antibodies for enzymes, venoms, poisonous vegetable proteins (ricin, abrin, etc.) and various bacterial hematotoxins. It is to be noted that these active substances are all similar to one another in being classed as large colloidal aggregates, resembling proteins, but not yet identified as proteins. The other group of immunity reactions is concerned with defense against foreign proteins, whether toxic or non-toxic and whether in solution or aggregated into cells (bacteria, corpuscles, tissue cells). In all the reactions of this group we deal with processes that tend to alter the colloidal state of the foreign proteins, by making them larger aggregates (precipitation, agglutination) or smaller aggregates (proteolysis, hemolysis, bacteriolysis, cytolysis), and in each case the reaction consists of two separable steps, sensitization and reaction. It is tempting to accept the view, championed especially by Friedberger, that in this second group of reactions but one and the same antibody is concerned, and that all the reactions are essentially the same, differing merely in the method by which the reaction is demonstrated. There are, indeed, many facts capable of interpretation as supporting this hypothesis, but there are still other items that do not harmonize with it, notably the lack of constant quantitative relations between the different reactions produced by the same immune serum, and as yet it is neither established nor completely disproved.

Anaphylaxis obviously comes into the second class of "proteolytic" immune reactions, and the identity of the sensitizing antibody and the precipitin has been strongly supported. Several observers have found that the relative proportion of precipitin and sensitizing antibody runs parallel in immune sera (Friedberger, Doerr and Russ) and I have

myself observed, in studying the development of antibodies in animals immunized to vegetable proteins, that the precipitins appear together with the capacity to confer passive sensitization (20). Richard Weil (21) added much support to this contention by finding that the precipitate obtained when immune serum reacted with the specific protein, and washed free from the serum, conferred passive sensitization to the specific antigen when injected into guinea pigs, if the proportion of antiserum and antigen had been suitably selected in producing the precipitate. Mechanical carrying down of serum antibodies was excluded by appropriate experiments. Although it was found that if immune serum is heated at 70° for one-half hour it loses its precipitating power but not its capacity for passive sensitization, Weil did not consider this as evidence of a separation of precipitin and sensitizin, but interpreted it as indicating that only the thermostable combining (haptophore) group of the antibody is essential to passive sensitization. "The fact that the coexistence of antigen in the same fluid may inhibit the precipitating power of antibody, while only partially interfering with the sensitizing function, may explain the divergence in the literature between those who maintain that precipitin and sensitizin run parallel in immune serum and those who deny this relationship." The precipitates produced in the precipitin reaction contain both antigen and antibody, since they sensitize both actively and passively, and Weil believed that both antigen and precipitin can be separately extracted from such precipitates by suitable procedures. Although Coca (22) was unable to repeat Weil's experiment with washed precipitin-antigen precipitate, yet he did secure further evidence that precipitin and sensitizing antibody are identical.

Granting the identity of precipitin and sensitizing antibody, one must consider the relation of both to the complement fixing antibody, which numerous observers have maintained is identical with the precipitin (23). The assumption of identity of the sensitizin and the complement fixation antibody is the basis of Friedberger's classical experiments on the production of a poison, supposed by him to be responsible for anaphylaxis, by digestion of sensitized proteins with complement. This topic will receive consideration later.

THE SITE OF THE REACTION

The earlier view that anaphylaxis resulted from reaction of antigen and antibody in the circulating blood was shaken by the observation that in passive sensitization of guinea pigs there is a latent period of at

least four hours before the animal becomes sensitive, and that simultaneous injection of antigen and antiserum produce no effects. This fact is difficult to explain if antigen and antibody react in the circulating blood, but fits in well with the hypothesis that the latent period is required for the antibody to enter the cells and render them sensitive. The humoral reaction hypothesis was rendered untenable as a complete statement by the observation of Schultz that isolated segments of intestine from sensitized animals reacted to the specific antigen by contraction, and by the development and refinement of this principle especially by Dale and Weil. It has now been established that isolated non-striated muscle tissue, washed as free as possible from blood, will give strong contractions immediately it is brought into contact with even very dilute solutions of the specific antigen, and only with the specific antigen. The possibility that even the most careful washing out by perfusion methods may leave blood in some of the capillaries, especially emphasized by Larson and Bell (24), seems to be eliminated as an explanation of the positive results, both by quantitative considerations and by the observation of Weil (25) that uterine muscle of a passively sensitized guinea pig does not react to the antigen during the latent period (even if not perfused and full of blood) although at this time the blood in the uterus contains an abundance of antibodies. On the other hand a very much smaller dose of antiserum, even when followed by careful perfusion twenty-four hours later, has then rendered the uterine muscle highly reactive to the specific antigen.

Besredka (1), who was one of the earliest supporters of the "cellular theory" of anaphylaxis, believed that the reaction occurred in the cells of the central nervous system, but this localization is excluded by many experiments which show that anaphylactic reactions occur in tissues entirely disconnected from the central nervous system, as well as by the muscle strip experiments mentioned previously. Another form of evidence of the non-humoral nature of anaphylaxis is furnished by transfusion experiments (Manwaring, Pearce and Eisenbrey, Weil and others) which show that the blood of a sensitized animal can be replaced with the blood from a non-sensitized animal without impairing its sensitivity. This form of experiment is particularly convincing when passive sensitization is employed with serum of known sensitizing power, since then it is possible to know surely that the transfused animal cannot have a sensitizing dose of antibodies left in its circulating blood (Coca, 26, von Fenyvessy and Freund, 27). Furthermore, Doerr and Pick have observed that even after all demonstrable anti-

bodies have disappeared from the circulating blood in the rabbit, fatal anaphylactic shock may be produced.

Not only do we have all this evidence of the cellular site of the anaphylactic reaction, but in addition there is abundant evidence that the presence of free antibodies in the blood interferes with the anaphylactic reaction (Weil, 28) presumably by binding the antigen before it can reach the "sessile" receptors within the cells. This is the explanation of the fact that repeated injection of guinea pigs with the antigen renders them immune rather than hypersensitive, and that a single large dose of antigen sensitizes less successfully than a small dose of the same antigen.

It may safely be said that at least the essential features of acute anaphylactic shock depend upon intracellular reaction between antigen and antibodies that have been withdrawn from the circulating blood and fixed within the cells. Whether the reaction which presumably occurs between antigen and the free antibodies in the circulating blood also plays any part whatever in anaphylaxis, is still undetermined. Zinsser (1) believes that the "radical cellular interpretation of all phases of the phenomena of anaphylaxis goes too far." Several observers have reported that when suitable quantities of antigen and antibody are used, it is sometimes possible to secure passive sensitization without the usual period of incubation; indeed certain workers have stated that simultaneous injection of antigen and antiserum may give rise to immediate anaphylactic reaction, especially in rabbits. The chief objection to these experiments is that the supposed positive results are obtained only occasionally, are usually very slight in character and without the necessary regularity to be convincing evidence of a fundamental principle, and Weil doubted the true anaphylactic character of such reactions as have been observed. Friedemann obtained the best results with simultaneous intravenous injections of antigen and antiserum in the proportions which were found to give the best precipitin reactions in vitro, which suggests the possibility that the symptoms observed in the injected rabbits were not those of true anaphylaxis, but anaphylactoid phenomena produced by capillary occlusion from the precipitate, or possibly from agglutination of red cells or fibrin formation incited by the precipitate, for it has been shown by Doerr that intravenous injection of antigen-antibody precipitates produces severe symptoms.

On the other hand, there is no doubt that antigen-antibody reactions do produce, at least in vitro, substances that are eminently injurious,

especially on intravascular injection, and it seems reasonable indeed to believe that such substances may be produced in the typical anaphylaxis reaction and play some part in it, even if we grant that the typical acute anaphylactic shock depends on reaction within certain tissue cells. For example, Manwaring and Kusama (29) found that although isolated lungs from sensitized guinea pigs exhibited strong bronchial constriction when perfused with the specific antigen, they gave still stronger reactions when the blood of the sensitized animal was present in the perfusion fluid. Some authors have advanced the hypothesis that the more chronic manifestations of anaphylaxis (such as fall of temperature in guinea pigs, the splanchnic engorgement and hemorrhagic enteritis in dogs) depend on humoral antigen-antibody reaction. Weil's demonstration (25) that during the latent period immediately following injection of immune serum, despite the presence of abundant antibodies in the blood, no anaphylaxis reaction can be obtained, either acute or chronic, makes difficult of acceptance the hypothesis that humoral reactions play any important part in anaphylaxis.

"ANAPHYLATOXIN" FORMATION AND ITS RELATION TO ANAPHYLAXIS

For a time in the history of the study of the anaphylaxis reaction the cause of the intoxication clearly seemed to lie in the products of cleavage of foreign proteins into toxic elements by the enzymes of the blood or the tissues. This "anaphylatoxin" hypothesis fitted most of the known facts so well, and was so perfectly logical, that it seemed almost inevitable, yet at the present time it appears to be untenable, in the face of existing evidence, as the final explanation of anaphylaxis. It rested fundamentally upon three sets of observations. First, the demonstration by Vaughan that from almost any protein a highly poisonous substance can be produced by cleavage under suitable conditions, and that this substance produces effects in animals much resembling the phenomena of anaphylactic shock or subacute anaphylactic intoxication. Second, the similarity of the effects produced by intravenous injection of Witte's "peptone" and other products of protein cleavage, to those of anaphylaxis. Third, the demonstration by Friedemann and Friedberger of the highly poisonous character of the products of reaction in vitro of antigen, antibody and complement. Taken together these observations seemed to indicate plainly enough that anaphylaxis was a typical case of poisoning by products of specific

proteolysis of antigens by serum complement after sensitization of the antigen by the specific antibody. Friedberger especially developed this hypothesis, although others had preceded him in producing intoxication of animals by injecting the products of antigen-antibody-complement reaction.

Friedberger showed that such toxic products were formed when precipitates, resulting from the reaction of antigen and specific precipitin (or bacteria or other cells that had been sensitized with the specific cytolytic amboceptor), were digested for a suitable period with serum complement. Also that the toxic product produced symptoms resembling those of anaphylaxis when injected intravenously into animals, and that these symptoms were the same irrespective of the nature of the antigen used, which is one of the characteristic features of anaphylaxis as well as of Vaughan's products of protein cleavage. Furthermore, too long digestion, or excessive sensitization, failed to produce the "anaphylatoxin," which supported the view that the poison represented an early stage in protein cleavage, which was destroyed by further digestion.

It was soon found, however, that this so-called anaphylatoxin formation was not a specific result of antigen-amboceptor-complement reaction, as identical toxic effects are produced with normal serum that has been digested with complement in the presence of absolutely inert materials, such as kaolin, barium sulphate and agar. Jobling and Peterson found that merely extracting the lipoids from normal serum and then letting it digest itself resulted in the same type of toxicity, and they agree with Bronfenbrenner (30) in attributing all the positive results in this type of experiment to a removal or inhibition of antienzymes present in normal serum, so that proteolysis then sets in. Even the clotting of blood or plasma may produce changes which render it highly toxic, even for the animal from which it came, with the same anaphylaxis-like manifestations that are produced by the toxic sera in the typical anaphylatoxin experiment.

Despite the undeniable fact that blood serum may be made highly toxic when digested with complement under various conditions, and that the symptoms which follow intravascular injection of such toxic sera resemble those of anaphylactic shock in many particulars, nevertheless the responsibility of such a production of toxic products of protein cleavage for the explanation of anaphylactic shock fails of proof. Some of the points that have been advanced against this so-called humoral anaphylatoxin theory are the following:

1. It does not fit with the latent period of passive sensitization. However, intracellular formation of anaphylatoxin might account for this phenomenon.

2. Complement is not essential, since animals deprived of free complement in the circulating blood may still give anaphylactic reactions. Here again, one may suggest the presence of intracellular or reserve complement.

3. All attempts to prove that complement is a proteolytic ferment have so far failed.

4. "Anaphylatoxin" activity has been produced in serum by digestion in the absence of complement, in the absence of antigen, and in the absence of antibody. On the other hand, if antigen and the specific antibody are injected simultaneously into the opposite jugular veins of a guinea pig, the animal shows no evidence of intoxication.

5. In the anaphylatoxin experiments the existence of capillary embolism or endothelial intoxication has not been excluded, and there is reason to believe that most of the observed symptoms are anaphylactoid rather than anaphylactic.

6. All attempts to demonstrate that the blood of animals in anaphylactic shock contains a poison responsible for the observed manifestation, have failed (see Weil, 31).

The theory of the dependence of anaphylactic intoxication upon the formation of non-specific poisons from proteins by some sort of ferment reaction, received its most severe blow in the clear-cut experiments of Dale (32) with the guinea pig uterus. If we grant, as it seems we must, that stimulation of nonstriated muscle is an essential, and probably the essential feature of acute anaphylactic shock, the Schultz-Dale method portrays the reaction without the possible invalidating effects of capillary thrombosis or agglutination emboli. The fact that, as Dale points out, the specific antigen, when introduced into a bath containing the muscle strip, produces a reaction *immediately it comes in contact with the muscle*, and quite as promptly as a preformed, diffusible stimulant, such as histamine, pilocarpine or pituitary extract, is difficult to reconcile with the theory that the stimulating agent is liberated through proteolysis. Moreover, says Dale,

> Apart from the absence of a coarse latent period, the rest of the time relations of the effect are very difficult to harmonize with any form of ferment theory. On such a theory, one would expect a gradual onset and slow rise to a maximum. But the onset is sudden, and the rate of progress to the maximum, when sensitiveness is fully developed and the dosage not too minute, is apparently limited

only by the contraction rate of the plain muscle. After a few minutes at the maximum the effect begins to decline, and the muscle is then insensitive to further doses. In the form of the contraction curve, as in the brief latency, the effect recalls that of a powerful stimulant drug added directly to the bath.

These observations seem to exclude not only the humoral site of anaphylaxis reaction, but also to rule out the possibility of even an intracellular digestion of a sensitized antigen by a ferment, whether serum complement or intracellular proteolytic enzymes.

But we cannot escape the fact that the manifestations of anaphylactic shock resemble in all respects those of an acute intoxication; furthermore, that histamine, the substance which produces the picture most closely resembling that of typical anaphylaxis, is a product of protein cleavage (see Abel and Kubota, 33; Dale, 34; Hanke and Koessler, 96). Not only does histamine cause bronchial spasm in guinea pigs, obstruction to pulmonary circulation in rabbits and fall of blood pressure in dogs, but applied to the skin or mucous membranes, it causes marked local urticaria resembling closely that of skin reactions in sensitized persons, and it does all these things in extremely minute dosage, comparable with the dosage of proteins used in anaphylactic reactions. Furthermore, its antecedent amino acid, histidine, is present in every known complete protein. Other pure chemical products of protein cleavage, such as methyl guanidine, have more or less similar effects. The chief respects in which histamine fails to account for all the phenomena of anaphylaxis are:

1. It fails to desensitize sensitized animals or tissues yet produces strong reactions in the uterus strip that has been thoroughly desensitized (Dale, 32).

2. Histamine does not produce the temperature reactions usual in anaphylaxis.

3. Histamine does not produce the changes in coagulability of the blood usual in anaphylaxis.

4. Quinine augments the susceptibility of sensitized animals to the foreign protein but does not affect the intoxication produced by histamine (M. I. Smith, 35).

Nevertheless, histamine seems to have a synergetic relation to anaphylactic shock, probably because the points of attack of histamine and anaphylactic reactions are identical (35). In respect to item one of the above list, we would not expect histamine to desensitize if it is the active product of an antigen-antibody reaction, since it is this reaction alone that is prevented by desensitization. As to items three

and four, it may well be that these phenomena result from other products of the antigen-antibody reaction, since presumably many different substances are produced. The effect of quinine (item four) might be found to be an influence on the antigen-antibody reaction if this possibility were investigated.

Moreover, in the so-called Abderhalden reaction we have evidence that protein cleavage does occur in mixtures of immune serum and specific protein, and that in this process highly active toxic substances are produced (Bronfenbrenner, 30). For whatever may be said concerning the specificity of this reaction, there undoubtedly does commonly occur a greater amount of proteolysis in such mixtures with the specific antigen than if some other protein is present. A careful study of the Abderhalden reaction (dialysis method) in my laboratory by Elsesser (36), using Osborne's purified vegetable proteins, showed that in spite of many atypical, irregular and illogical results, "there is an obvious tendency for a substrate to react more often and yield stronger reactions when tested against its homologous immune serum, than when tested against a heterologous immune serum."

Also, we cannot afford to overlook the important fact that racemized proteins, which are characterized by being incapable of attack by enzymes in vitro, or of being digested and metabolized in vivo, are also incapable of serving as antigens, although derived from proteins which in the original state are highly antigenic.

It being established that immune sera do cause proteolysis of specific antigens, or at least produce diffusible cleavage products when in contact therewith, and since the intermediate products of proteolysis are unquestionably toxic, how can we discard entirely the possibility that such substances do arise, and play at least some part, in anaphylaxis, even if we can exclude proteolysis as the responsible factor in the immediate contraction of non-striated muscle both in vivo and in vitro? Furthermore, since the so-called anaphylatoxins are so readily produced, and exert such a profound physiological effect, their importance in human pathology must be considerable, even if they are not responsible for typical acute anaphylactic shock. Such observations as those of Quenu and Delbet, Cannon (37) and others, on the relation of disintegration of traumatized muscle tissue to traumatic shock, come to mind in this connection.

Although we are justified in defining sharply the limits of what shall be considered as true anaphylaxis in discussing and studying that topic, phenomena which we must exclude as outside those limits are not

therefore any less important. They merely are other problems, and presumably related problems, but how much bearing they have on specific anaphylaxis, if any, we do not know. From this open viewpoint as a contribution to our knowledge of serum toxicity, the extensive studies of Novy and his collaborators (38) are of interest. They corroborate and extend the findings of others, that the serum of a guinea pig may be made toxic for guinea pigs after incubating with agar, inulin, or small formed particles, such as trypanosomes or inorganic precipitates, so that from 1 to 3 cc. may be a lethal intravenous dose. Rat serum, which is toxic to guinea pigs even when untreated (3 to 4 cc. sometimes being fatal) may have its toxicity enhanced until 0.25 cc. is lethal. Also, agar and other substances which render serum poisonous in vitro are also very toxic when injected intravenously, as little as 9 mgm. of agar per kilo of guinea pig being fatal. But even unformed substances may render serum toxic, e.g., peptone solutions, and distilled water is said to produce "anaphylatoxin" in rat serum, which Sachs (39) would explain by the precipitation of the serum globulins by the distilled water, thus disrupting the colloidal equilibrium and permitting the formation of toxic forms of colloids. But the proof that such phenomena are the same as those underlying the specific antigen-antibody reaction of anaphylaxis has not been furnished, and the term anaphylatoxin is misleading, especially as we have as yet no proof that a toxic material is formed in and responsible for true anaphylaxis reactions. Moreover, the "anaphylatoxins" thus formed (better called "serotoxins") seem to have little effect on rabbits, although these animals are far from insusceptible to anaphylaxis. Nor can we reconcile the observation that the mere coagulation of blood renders the serum highly toxic, with the known facts of anaphylaxis. Unquestionably the readiness with which blood becomes poisonous is an important phenomenon, and presumably it plays an important part in pathology, but that this phenomenon is responsible for, or even related to, anaphylaxis, has yet to be demonstrated. Novy and DeKruif make this statement: "It will be seen that the phenomenon of anaphylatoxin production, which includes anaphylaxis, is not connected with the problem of immunity except so far as it shows the ready response of a labile blood constituent to an alien substance." But anaphylaxis *is* a definite immunological phenomenon, and if the phrase "which includes anaphylaxis" is omitted from this statement most immunologists will agree, accepting the independence of so-called anaphylatoxin formation from immunological reactions in general and from anaphylaxis in particular.

But if we discard the anaphylatoxin theory of anaphylaxis, we are left without an explanation of the very striking phenomena of anaphylactic shock, for no satisfactory substitute hypothesis has been proposed. Dale (32) suggests that "A disturbance of the conditions of colloidal solution is set up in the muscle fibre," which is not altogether satisfactory, although it may well be true. Weil speaks of a "cellular discharge" which is even more vague. Danyz (40) has proposed the hypothesis that anaphylaxis is merely an intracellular or intravascular "indigestion," because of the impossibility of transforming colloidal antigen into crystalloids, but this does not fit at all the known facts of the character of the reaction in isolated tissues. None the less, the fact that slight disturbances in the equilibrium of plasma colloids render them highly toxic is an outstanding fact, and it becomes easily understandable that a similar alteration in colloidal equilibrium within the cell protoplasm may produce equally profound intoxication of the cell, and at present there is a growing tendency to seek an explanation of anaphylaxis in the domain of colloidal chemistry (see Zinsser, 1, Sachs, 39).

PATHOLOGICAL PHYSIOLOGY

Anaphylactic reactions have been produced in many species of mammals, but only in the guinea pig, rabbit and dog has the process been thoroughly studied. The striking fact soon developed that in each species the phenomena were different, and always the same in that species without relation to what antigen was being used. No matter how widely separated in biological origin or chemical character the antigenic protein might be, the symptoms produced were always the same, a fact which gave support to the hypothesis that the poison responsible was derived from the proteins of the sensitized animal itself. Auer and Lewis demonstrated that in the guinea pig the typical respiratory difficulty was caused by a spasm of the bronchial musculature. Anatomical studies showed that the guinea pig has an exceptionally highly developed musculature in its bronchi, the finer bronchioles being practically nothing but muscular tubes, possessing also a thick folded mucous membrane. More recent experiments have shown that this spasm of the bronchioles is not accompanied by a corresponding spasm in the blood vessels, since perfusion fluids pass through the distended, anaphylactic guinea pig lung without obstruction.

In the rabbit this lung inflation is not observed, or at least not to any corresponding degree, the symptoms being chiefly circulatory.

There is a fall in arterial blood pressure, slowing of the heart beat, with marked distention of the right side of the heart. Auer believed that these changes resulted from a direct effect on the muscle of the right ventricle, and the isolated hearts of both rabbits and guinea pigs do give reactions in vitro. Coca (41) found that during anaphylactic shock the rabbit's lung develops a greatly increased resistance to the passage of perfusion fluids, with embolism and thrombosis excluded, apparently from a spasm of the muscular coat of the arterioles, which explains the observed symptomatology. Here again histological study offers an explanation, since the pulmonary arteries of the rabbit present a remarkable degree of muscular development, analogous to that of the guinea pig bronchioles.

The symptoms produced in the dog are again quite different and distinctive. Here the liver and splanchnic circulation show the most marked effects, becoming tremendously engorged with blood, so that the systemic blood pressure falls because depleted of blood. Weil (42) showed that this was a local reaction, for when he injected the antigen into one of the branches of the portal vein supplying a single liver lobe, the congestion was limited to this lobe. This hepatic effect has generally been attributed to an action on the liver cells, especially since it is accompanied by alterations in the coagulability of the blood, and other evidences of altered hepatic function, such as reduced capacity to form urea from ammonium lactate (Rumpf, 43). A more plausible and logical explanation has been advanced by Simonds (44) who states that the walls of the hepatic veins of the dog differ from those of all other animals in the great development of their musculature, inferring that the hepatic and splanchnic congestion are the result of spasm of the hepatic veins. As yet the necessary tests to show that occlusion of the hepatic veins does occur, and that this of itself can account for the observed phenomena in canine anaphylaxis, have not been reported.

If this explanation is correct, the characteristic features of acute anaphylactic shock in different species depend merely on differences in the distribution of non-striated muscle in the different species. We have abundant experimental evidence that unstriated muscle of all tissues or organs is stimulated to contraction in the anaphylactic reaction, which explains the micturition and defecation, and perhaps the erection of the hairs and the exophthalmos observed in anaphylactic shock. Whether the marked changes in the heart action, sometimes in the nature of heart block, revealed in the anaphylactic rabbit

by electrocardiographic methods (Auer and Robinson, 45) depend on spasm of the coronary arteries, on reaction of the heart muscle itself, or on the pulmonary obstruction, is not known. Presumably the stimulation of unstriped muscle tissue is universal [thus, in the dog severe bronchial constriction does occur even with the liver and other abdominal viscera excluded from the circulation (Pelz and Jackson, 46)], and the symptoms observed depend upon the degree of development of these muscles, or their strategic location, in different organs of different species.[2] Corroborating this view is the fact that fatal anaphylaxis reactions in man are practically limited to persons suffering from chronic asthma.[3] In such persons there is a profound thickening of the bronchial wall, together with great hypertrophy of the bronchial musculature, which facts are brought out strikingly in the study of the pathological anatomy of bronchial asthma in this laboratory by Huber and Koessler (51). Such asthmatic persons, having bronchi like those of the guinea pig, react in the same way as the guinea pig when injected with the specific antigen. On the other hand, it may occasionally happen that the manifestations may affect chiefly the splanchnic area, as in the dog (see case reported by Gurd and Roberts, 52), but the reason for this localization in these cases has not been sought in the light of Simond's observations.[4]

[2] The theory that leucocytic blocking of pulmonary capillaries is important (Werigo, 47) is not supported by many studies of the lungs in anaphylaxis, and is contradicted by the lack of any decrease in severity of reaction in guinea pigs with greatly reduced number of leucocytes from Thorium-X treatment (Corper, 48).

[3] A remarkable instance of fatal anaphylaxis in man has been reported by Jeanbrau (49). The patient had been given antitetanus serum, and three weeks later received a transfusion of blood from a man who had received a dose of antitetanus serum that very morning, death resulting in a few seconds. On the whole, however, man does not seem to be highly susceptible to anaphylaxis, at least as compared with the guinea pig; thus, Böttner (50) calls attention to the fact that collargol has had a widespread use by intravenous injection at intervals suitable to produce sensitization and shock, without recorded serious results, although it contains 25 per cent of protein and produces reactions readily in guinea pigs.

[4] Since examining Doctor Simond's preparations of the dog hepatic veins, I have been interested in the great muscular development of the human adrenal veins. Although I have not yet had an opportunity to make measurements of this material, I have often noticed the large amount of muscular tissue in the central veins of the adrenal, apparently greater than that in other veins of corresponding caliber in man. In view of the relation of epinephrin to anaphylactic shock this observation, if corroborated by actual measurements, may be of some significance.

There is no pathological anatomy of anaphylactic shock that cannot apparently be explained by such a mechanism as that described above. The anaphylactic poison, if there is one, would seem to be after the order of the alkaloidal poisons, since it produces immediate but transitory effects, without evident structural alterations in the tissues. Such anatomical alterations as are observed, such as hemorrhages or waxy degeneration of the voluntary muscles of respiration (53), are ascribable to the asphyxia and circulatory obstruction. Longcope (54) and Boughton (55) have observed lesions in the kidneys, hearts and livers of rabbits and guinea pigs subjected to repeated anaphylactic shock, as well as to repeated protein injections without shock, but in view of the frequency of spontaneous lesions in animals kept under similar conditions and subjected to much manipulation, it cannot be said positively that the lesions described resulted from anaphylactic shock, or from foreign proteins (see Bell and Hartzell, 56). Nevertheless, we have in the Arthus phenomenon a striking proof that local tissue injury may result from anaphylactic reactions, whether through direct cellular injury or protracted vascular occlusion, and so it would seem eminently probable that tissue injury may result elsewhere than in the skin under similar conditions.

Weil (57) has described changes in the liver of dogs dying within one hour of anaphylactic shock, consisting of a cloudy swelling and in one case disseminated necrosis, but it is possible that these changes may be the result of the intense asphyxia from which the liver suffers. It is difficult to understand how injury to the liver cells themselves could account for the engorgement, as Weil suggests, and equally difficult to accept is Dale's idea of a selective toxic action on the endothelial cells of the liver. Presumably the decreased coagulability of the blood depends on injury to the liver cells and release of anticoagulating substances,[5] for Weil found that blood containing the specific antigen loses its coagulability when perfused through the liver of a sensitized dog, and poisons which injure the liver produce a similar decrease in coagulability. Also, autolysis of the liver is said to be hastened in sensitized animals (95).

Apparently some change takes place in antigens which enter the anaphylactic liver, for Manwaring and Crowe (59) found that antigen perfused through the liver of a sensitized guinea pig lost its toxicity

[5] According to Bulger (58), however, the increase in antithrombin is not great, and he ascribes the observed changes in coagulability to variations in thromboplastin, together with a marked increase in fibrinolysis.

for sensitized guinea pig lungs, although normal livers did not have this effect, and Falls (60) observed that larger doses of antigen are required to produce shock when injected into the portal vein than when introduced into the jugular. Weil (31) has found evidence that antibody and antigen coexist within the reacting tissues for some time, the antigen being gradually destroyed, a process which may perhaps be accompanied by injury to the cells, but L. Loeb (61) could not find that thyroid or uterus from sensitized guinea pigs transplanted into other animals were modified in their vitality as tissue grafts, by injection of the specific antigen into the host.

Although we recognize the systemic anaphylactic reaction only as it affects non-striated muscle tissue, it may well be that this is merely because such tissue alone registers the effects in a graphic manner. We do not know whether any or all other cells of the sensitized animal are affected when they come in contact with the specific antigen. Still less do we know what happens within the sensitized muscle cell when the antigen and antibody meet. Dale has shown that the response is immediate, as when a potent diffusible drug acts on a cell, but whether the reaction changes the colloidal state of the cell, as has been vaguely suggested, or whether it causes a discharge of energy after the order of an electrical stimulus, or something entirely different, are matters concerning which we have not the slightest information.

Metabolism studies may show an increased toxicogenic destruction of protein (62), but the increase in amino acids, presumably resulting from proteolysis within the tissues of the reacting individual, is not large enough, if it occurs at all, to be significant. Thus, Auer and Van Slyke (63) were unable to find a demonstrable increase in free amino nitrogen in the anaphylactic guinea pig lung, nor do the livers of guinea pigs dying in anaphylactic shock contain appreciably more non-coagulable nitrogen than normal livers (64). However, in anaphylaxis in guinea pigs, as well as after peptone poisoning, there is a considerable increase in non-coagulable and urea nitrogen in the blood, as well as a slight increase in amino nitrogen, but it is not known whether this comes from the tissues or from the antigen-antibody reaction in the blood, although the former seems more probable (see Hisanobu, 65). Manwaring (66) found that no appreciable loss of the antigenic protein occurred during its perfusion through the organs of sensitized animals.

DESENSITIZATION AND ANTI-ANAPHYLAXIS

An important and characteristic feature of the anaphylactic reaction is the refractory condition which the recovered animal exhibits immediately thereafter, provided that the amount of antigen injected has been sufficiently large, and isolated muscle strips exhibit the same phenomenon. If the reaction is produced by minimal quantities of antigen, the animal may still react to a second dose of the same antigen. Evidently this condition of desensitization depends upon an exhaustion or saturation of the antibodies. The desensitization is quite as specific as the acute reaction itself, for if the guinea pig has been sensitized to two different proteins and recovers from the shock induced by one of them to which it thus becomes entirely refractory, it will still react to the second protein. I have had much occasion to take advantage of this phenomenon in determining the purity of protein preparations, and in studying the relations of such proteins, and have found it to exhibit a reliable specificity. It was noted early in this work, however, that the reaction to the second protein is seldom as severe as it would be if it were a primary reaction. This has been repeatedly confirmed by others, and Dale (32) found that the isolated uterus from an animal sensitized to more than one antigen showed a similar decrease in reactivity after a maximal reaction to one of these antigens. Presumably this is due at least in part to exhaustion of the reacting mechanism, and perhaps in part, in the case of intraperitoneal injections, to an impaired absorption because of inflammatory reactions and obstructed lymphatics.

Desensitization can be accomplished without anaphylactic shock if the antigen is introduced slowly enough, which is usually the case when the injection is subcutaneous. Conversely, the production of observable anaphylactic shock requires that the toxic dose of antigen reach the antibodies within a brief period.

Weil (67) made a quantitative study of desensitization following passive sensitization with standardized amounts of antibody, and found that the quantity of antigen needed to produce a reaction in a partially desensitized pig did not bear a constant relation to the amount of unsaturated antibody presumably left after the first reaction, as shown by the following table:

Sensitization: 0.3 cc. immune serum intraperitoneally

GUINEA PIG SERIES	DESENSITIZING DOSE OF HORSE SERUM	MINIMAL ANAPHYLACTIC DOSE OF HORSE SERUM, INTRAPERITONEALLY
	cc.	cc.
1	0.005	0.1
2	0.01	0.5
3	0.02	0.5
4	0.05	0.5

As the minimum anaphylactic dose of antigen after partial desensitization is much larger than before partial desensitization, it would seem that the reactivity or affinity of the antibody had been much reduced in the desensitizing process. This alteration in quantitative relationship between antigen and antibody is distinctly different from the fixed quantitative relations observed between precipitinogen and precipitin, and Weil pointed out the resemblance to the Danyz phenomenon observed in toxin-antitoxin neutralization; i.e., the changed quantitative relationship produced by adding antitoxin to toxin in small fractions instead of a single neutralizing quantity. As he stated,

> The same antibody, when in solution as precipitin, combines with antigen quantitatively, and in strictly constant proportions, to form precipitate; when attached to the cell, as sensitizing antibody, it combines with antigen in varying proportions. The living cell, therefore, modifies its properties.

Coca (1), however, makes the valid suggestion that these observations can be explained by the assumption that the sensitizing serum contains antibodies of varying avidity for antigen, whereby those with the greatest avidity are saturated with the first partial desensitizing dose, leaving the less reactive antibodies to unite with the second dose of antitoxin. These suggestive and important quantitative studies have not been extended since the untimely death of Doctor Weil removed one of the most brilliant and productive students of anaphylaxis.

Sensitized animals may be refractory to anaphylactic reactions not only because of saturation or exhaustion of the fixed antibodies, but because the presence of free antibodies in the circulating blood prevents the antigen from coming in contact with the intracellular antibodies, i.e., the condition of anti-anaphylaxis. There has been much confusion in the literature because of the careless use of the term "anti-anaphylaxis" to cover all refractory conditions. This term should logically be applied only to a resistance due to antibodies, using the

term desensitization for the condition which it describes, as indicated in previous paragraphs.

It is this form of antibody resistance, true specific anti-anaphylaxis, which accounts for the failure to obtain anaphylactic shock in guinea pigs immunized to a foreign protein by repeated injections. This assumption, which was made early in anaphylaxis work, seems to have been converted into a certainty by the demonstration of Manwaring and Kusama (29) that the isolated organs of such immunized guinea pigs, washed free of circulating antibodies, are very highly sensitized to the specific antigen. Weil (68) further showed that if small doses of antigen are injected subcutaneously into passively sensitized guinea pigs that have been given intravenous injections of large protective doses of antiserum, the uterus is not desensitized as it otherwise would be, since the antigens are kept from uniting with the intracellular antibodies.

Antisensitization is a somewhat similar phenomenon described by Weil (69). If a guinea pig is given a single dose of rabbit serum several days before a sensitizing dose of serum from a rabbit immune to a foreign protein, the usual passive sensitization does not take place. This is explained by the development in the guinea pig of antibodies to the rabbit serum, which protect the guinea pig's tissues from the antibodies of the immune rabbit serum. In proof of this conclusion is the fact that such preliminary injection with rabbit serum does not prevent passive sensitization with the serum of a guinea pig immunized to foreign protein. Also, Manwaring and Kusama (29) showed that in immunized guinea pigs the bronchial musculature is highly sensitized but that the presence of the blood of the immune animal will prevent reactions with the specific antigen in lung perfusion experiments.

Another form of refractoriness, which is essentially a desensitization, may occur when a large sensitizing dose is given. Differences may be observed even when the larger sensitizing doses are not very large; thus Thomsen (70) states that a sensitizing dose of 0.004 cc. of serum produces a maximum sensitization more quickly than 0.1 cc., although the final maximum degree of sensitization is the same with each dose. This seems explainable on the basis of the persistence of antigens in the blood during the time of formation of the antibodies and their fixation within the cells, whereby they are saturated with the antigen gradually as fixation takes place, without at any one time sufficient reaction taking place to be observable. Weil believed that antigen

and antibody might coexist in the cells without union, as they are known to do in the circulating blood. Of course, when a very large sensitizing dose of antigen is given there may be engendered sufficient antibodies to produce the condition of antianaphylaxis from the excess antibodies in the blood.

Still a different form of interference with sensitization has been described by Julian Lewis (71), who found that an amount of protein that will produce a marked anaphylactic sensitization when injected alone into a guinea pig, will fail to do so if injected together with, or twenty-four hours after, a much larger amount of another protein. Thus, in a series of guinea pigs sensitized with 0.1 cc. horse serum mixed with decreasing quantities of dog serum from 2 cc. to 0.001 cc., no reactions were obtained in the animals receiving from 1 to 2 cc. of horse serum, severe reactions but with recovery in those receiving from 0.25 to 0.01 cc. dog serum, as compared with promptly fatal reactions in the animals that received no dog serum.

The serum from a rabbit immune to horse serum, which will markedly sensitize a guinea pig to horse serum, will fail to do so if injected with, or twenty-four hours after, a large dose of another protein.

These results may be explained by the conception that the number of receptors in the body that can unite with a foreign protein is limited. The inhibiting protein, if present in large amount, combines with all, or almost all, of these receptors. Hence, another protein injected with it, or after it, is prevented from being combined in sufficient amount to stimulate the active production of antibodies. And when a large amount of protein is injected with or after a sensitizing dose of immune serum, the combination of the latter with the cell receptors, which is necessary for passive sensitization, is prevented in the same way.

I have found that the condition of antisensitization may be produced through feeding (72) as well as by injection, and presumably this condition explains a number of phenomena observed both in experimental animals and in man.

Guinea pigs bred from mothers fed with oats were, as soon as weaned, put upon a diet of egg albumen and carrots. Other young pigs from the same stock were raised upon oats and carrots. The latter animals, after reaching a weight of 250 to 300 grams, did not give anaphylactic reactions when injected with 0.05 gram of a protein obtained from raw oats, and if given small doses, such as ordinarily given for sensitizing, they were not rendered sensitive to subsequent injections of 0.05 gram of oat protein. Some of the pigs which were raised to 200 to 250 grams weight without oats were found to give a typical reaction of moderate severity when injected once with 0.05 gram oat protein, apparently from passive sensi-

tization conferred by the mother. Others gave no reaction. After the animals fed without oats were somewhat older, 350 to 400 grams, they reacted much less strongly or not at all to oat protein, as if this inherited passive sensitization were passing off, as passive sensitization normally does; such pigs, if given sensitizing doses of oat protein, are found to be sensitive to this protein three weeks later, giving well-defined reactions of moderate severity. Hence the conclusion seems warranted that if guinea pigs are raised on oat proteins they cannot be made to give anaphylactic reactions with oat proteins, but if raised without oats they may be sensitized to oat proteins, just as they can be to other proteins not usually in their food. These experiments support experience obtained previously with zein, that guinea pigs become immune to the chief vegetable proteins of their food.

Guinea pigs raised from the time of weaning on a diet of egg protein (Merck's dried egg albumen) and carrots were found to give strong anaphylaxis reactions when injected with egg albumen between the thirtieth and sixtieth days, but later they reacted less strongly, and after the one hundredth day of feeding they gave but slight reactions to 0.1 gram dried egg albumen. At this time sensitization with egg albumen can be obtained to only a slight degree, such guinea pigs given injections of egg albumen showing but slight reaction to a subsequent dose of egg albumen, while control pigs fed on oats and carrots gave severe, usually fatal, reactions to corresponding injections of egg albumen. Apparently, then, daily absorption of animal protein in the food at first renders guinea pigs hypersensitive to this protein, but if the feeding is kept up for a long enough period the animals become refractory to the food protein and are so immunized that they cannot be sensitized to this protein.

A series of guinea pigs which were raised on bread and cow milk for ten weeks were found at the end of this time to be still highly sensitive to milk, dying promptly when given 1 to 3 cc. of milk intraperitoneally. Apparently this length of feeding is not sufficient to render guinea pigs immune to milk. These results are not at all in harmony with those of Besredka (73), who did not succeed in sensitizing with milk by either oral or rectal administration, but did find that sensitized guinea pigs were made refractory to intracerebral injections of milk if previously given milk by either of these routes. They are corroborated, however, by the observation of Schlutz and Larson (74) that the blood of infants exhibiting the exudative diathesis when injected into guinea pigs renders them passively sensitive to milk.

Non-specific transitory reduction of the intensity of anaphylactic reactions may be produced by intravenous injection of numerous substances shortly before giving the intoxicating dose. Peptones, trypsin, various inorganic salts, urine and foreign proteins other than those used in sensitizing, are among these non-specific desensitizers, and their mechanism of action is not understood. Kopaczevski and Vahram (75) have found that sodium oleate has this effect to a marked degree, which they ascribe to a lowering of surface tension of the blood, on the hypothesis that anaphylactic shock depends upon a plugging of

capillaries by colloidal flocculation. Although perfusion experiments demonstrate that such gross capillary occlusion does not occur in the anaphylactic lung, yet the various items of evidence that changes in the colloidal state of the blood are of importance in the phenomena of "anaphylatoxin" intoxication make this observation of significance. The protective action of NaCl has usually been ascribed to its interference with complement action, but Richet (76) attributes it to its action on nerve cells, and others would ascribe it to the effect on the colloidal state of the cells and fluids. Thomsen (70) refers non-specific interference with the anaphylaxis reaction to a reduction in the speed of reaction between antigen and antibody, whereby a greater amount of antigen is found necessary to produce shock, but no explanation is offered as to why the speed of reaction is reduced.

THE SPECIFICITY OF THE ANAPHYLAXIS REACTION

If we limit the evidence on this subject, as we properly should, to that obtained in unquestionable typical anaphylaxis with protein antigens, it may be said that the anaphylaxis reaction exhibits exactly the same degree of specificity that we obtain with the complement fixation or precipitin reactions. In all cases group reactions are obtained when sera or extracts of tissues, etc., derived from closely related species, are used, and likewise, as emphasized by Magnus and others (20) and confirmed for the isolated uterus (by Dale, 32), increasing the antigenic doses tends to broaden the scope of the reaction.

Experiments with such hopelessly complex mixtures as are furnished by serum, organ extracts, or even egg whites and milk, cannot be expected to yield much valuable information concerning specificity. A common error in the literature is the statement that the anaphylaxis reaction, or some other immunological reaction, is not absolutely specific, since slighter reactions are obtained with sera of related species, etc. What is not specific in such experiments is the antigen used, not the reaction. If the antigens are isolated proteins, purified as completely as possible, striking illustrations of specificity may be demonstrated both by sensitizing and by desensitizing with specific proteins. For such work the proteins from seeds and nuts seem to be especially useful, for commonly they represent stored proteins, apparently set apart with a small number of other proteins as compared with the conditions in blood and tissues; furthermore, many of them are crystallizable, some are soluble in alcohol, and others present strik-

ing solubility characteristics, so that they are especially susceptible of separation and purification. Of the animal proteins those that resist heat coagulation, such as casein and ovomucoid, or those that crystallize, of which egg albumin is the chief example (for the crystallizable hemoglobins are not antigenic) are the best materials.

Working with such proteins, it has been possible to demonstrate conclusively that biological origin is not, *per se*, the thing that determines the immunological specificity of a given protein, but its chemical composition. The fact that the sera or tissues of different species can commonly be distinguished from one another by immunological reactions, indicates therefore that as a rule the proteins of different species are chemically different from the proteins of other species, even when these proteins have similar functions and exhibit close physical and chemical similarities, e.g., horse serum and ox serum. But there are many exceptions to this rule of immunological differences between the proteins of different species, e.g., goat casein and cow casein show to a considerable degree interchangeable immunological reactions, whereas they show a more distinct specificity against the serum proteins of the same species as furnished the casein. Thus, cow milk casein can be differentiated even from lactalbumin of the same species more sharply than from casein of goat's milk by complement fixation (Bauer), and cow casein and goat casein sensitize mutually to one another (72). A study of the proteins of hens' eggs by the anaphylaxis reaction showed that five distinguishable antigens could be differentiated from one another, and, furthermore, that these corresponded to five different proteins that had been differentiated by chemical methods, namely, ovomucoid, ovovitellin, albumin, and two different globulins (72).

Extended studies with milk proteins have also developed the same fundamental fact, that antigens which can be distinguished from one another by means of the anaphylaxis reaction are usually found to be readily distinguishable from one another by chemical means (77). The fact that different protein fractions of serum show stronger reactions when the same fraction is used as both sensitizing and intoxicating doses than when different fractions are used (Dale and Hartley, 78; Kato, 79) probably indicates a similar relationship.

The study of vegetable proteins also brings this out clearly, and furthermore indicates that we may have chemically distinct although similar proteins, which show corresponding distinctions and relations. Thus, hordein of barley is chemically similar to gliadin of wheat and rye, although showing sufficient chemical differences to establish it as a

distinct protein, but gliadin of wheat is chemically indistinguishable from gliadin of rye. Guinea pigs sensitized with gliadin from wheat or rye give strong anaphylactic reactions when injected with hordein of barley, but not as strong as with the homologous proteins, and the same results are obtained with injections in the reverse order. After such cross reactions, however, the recovered animals are not found to be desensitized to the homologous proteins, which indicates either that there are two or more individual proteins in the preparations of hordein and gliadin, one of which is common to both, or else the presence in gliadin and hordein of common and specific reactive groups. The chemical evidence is in favor of the latter conclusion. On the other hand, the chemically indistinguishable gliadins of rye and wheat interact against one another as if they were one and the same protein (7). Other examples of crossed reactions between unlike proteins have been observed; for example, legumin from pea renders guinea pigs as sensitive to legumin from vetch as it does to itself, and vice versa, and legumin from other seeds shows close immunological relation. On the other hand, a chemically distinguishable protein from another leguminous plant, glycinin from soy bean, does not sensitize to vetch legumin, or vice versa (80). The so-called natural proteoses of seeds show many striking illustrations of the same sort (11).

These and other experiments have led to the conclusion that

> the entire protein molecule is not necessarily involved in the specific character of the anaphylaxis reaction, but this is developed by certain groups contained therein, and that one and the same protein molecule may contain two or more such groups. It may well be that the intact protein molecule is involved in the reaction (for there is but little evidence that anything less than an intact protein molecule is capable of producing the typical reaction), but that certain groups determine the specificity (Wells and Osborne, 81).

Also, "since chemically similar proteins from seeds of different genera react anaphylactically with one another, while chemically dissimilar proteins from the same seed in many cases fail to do so, we must conclude that the specificity of the anaphylaxis reaction depends upon the chemical structure of the protein molecule" (80). Corroborative evidence has been furnished with precipitin reactions by Landsteiner and Lampl (82).

As to what groupings in the protein molecule determine the specificity, little is known. Obermayer and Pick (83) believed that the specificity characteristic of species was determined by the aromatic radicals, since iodizing or diazoting proteins removed this specificity

when tested by the precipitin reaction. Iodized proteins tested by the anaphylaxis reaction, however, I have not found to have altered in range of specificity from the original proteins (8).

That the anaphylaxis reaction is a less delicate means of determining specificity than the precipitin reaction, and especially the complement fixation reaction, is shown by the fact that when animals are immunized to hordein a serum of low antibody titer reacts only with hordein, but with antihordein serum of high titer both precipitin and complement fixation reactions are obtained with gliadin of wheat or of rye and passive sensitization is given with the same proteins (20).

Probably the more general use of the uterus strip method will enhance the delicacy of anaphylaxis work in the future, since this offers less uncontrollable factors than experiments with the intact animal. A good illustration of this is furnished by a recent report from Dakin and Dale (84). Previous workers have stated that the egg proteins of different species of birds are not readily distinguishable by anaphylaxis reactions, or at least that they show a less marked specificity than serum (for review, see 72). Welsh and Chapman found by the precipitin reaction, using the saturation test, that antiserum for any egg white contains a general avian antisubstance reacting with all sorts of bird egg proteins, and an independent antisubstance specific for the species. By the anaphylaxis reaction, however, such tests as have been made have shown a striking absence of specificity. Dale and Hartley found, in preliminary experiments, that crystallized albumins from hen eggs and from duck eggs sensitized mutually to one another even with the guinea pig uterus method, but by using quantitative methods and by checking up with the desensitizing test, it was later possible for Dale to detect differences between the two. The importance of this observation lies especially in the studies carried out by Dakin on the chemistry of these two albumins. When hydrolyzed, their amino acid content is very similar, but when the proteins are first racemized and then hydrolyzed it is found that there are distinct differences in the degree of racemization of the leucine, aspartic acid and histidine. There being reasons for believing that the amino acids which escape racemization occupy the terminal positions in the peptid chains of which the protein molecule is built, this result indicates that there are at least structural or spatial differences in these two proteins. In other words, a close relationship by immunological tests is here associated with chemical similarity, and a slight difference in chemical structure is found which presumably accounts for a slight

immunological difference that can be detected only by the most sensitive methods. This point of view is quite in harmony with that developed by Osborne and myself from our studies of the immunological specificity of the vegetable proteins. Although not based on the anaphylaxis reaction, but on precipitins, the corroborative results of Landsteiner and Lampe (82) may be recalled here; namely, that if one derivative (halogenized, azotized, etc.) of a protein reacts with another sort of derivative, the position in the molecule of the substituted radicals is identical or closely related. That is to say, cross reactions depend on chemical relationships, and the specificity is determined by relatively small portions of the large antigen molecule.

The observation that the location in the molecule of definite groups is indicated by their immunological reactions can best be explained as depending on spatial correspondence of antigen and antibody, just as Emil Fischer assumed for the specific action of ferments in his comparison to "lock and key." Here again we get evidence that both chemical composition and spatial relations are concerned in determining specificity. Presumably there are also innumerable isomeres that cannot be distinguished by our present methods, which correspond to the racial and individual differences which are so obvious and yet not to be detected by serum reactions.

Contemplating the possible number of variations in the arrangement of the amino-acids in a protein which the great number of these radicals provides, there is no difficulty in understanding the existence of an almost limitless number of specific distinctions between proteins. Abderhalden, indeed, calculated that the 20 amino acids we find in proteins could form at least 2,432,902,008,176,640,000 different compounds, and this without including possible compounds varying in quantitative relations. A contribution to the chemical basis of specificity has been made by Kossel (85), who finds certain relations in the proportions and groupings of the scanty number of amino-acids that make up the protamines and histones of sperm, to be characteristic of the sperm of certain species and families.

THE RELATION OF ANAPHYLAXIS TO SERUM DISEASE, TO DRUG HYPERSENSITIVENESS AND TO HAY FEVER

At present the evidence in respect to these topics is so contradictory as to defy positive conclusions. In respect to certain acute serum reactions, such as one in which a person with "horse asthma" dies in a few moments of bronchial spasm after receiving an injection of horse

serum, it seems evading the obvious to attempt to interpret the occurrence in any other way than as true anaphylactic shock, resulting from a specific antigen-antibody reaction; and as such a sensitized person usually exhibits also typical acute local reactions immediately after intracutaneous introduction of the most minute amounts of horse serum, or of other horse proteins as well, it seems difficult to deny, at least in such a case, that the cutaneous reaction of hypersensitiveness represents again a true specific antigen-antibody reaction, and is a true manifestation of anaphylaxis. Furthermore, it is sometimes possible to desensitize a person to the protein to which he is sensitive, removing both the systemic and cutaneous reactivity.

A particularly interesting demonstration by means of passive sensitization from man to man, of the true antigen-antibody nature of horse asthma and the specific skin reaction accompanying it, has been described by Ramirez (86). A man who had never had asthma, hay fever, urticaria, or any similar condition, received a transfusion of 600 cc. of blood from a man with typical horse asthma, who gave a cutaneous reaction to horse dandruff in 1–50,000 dilution. Two weeks later the transfused patient went for a carriage ride and within five minutes had a typical attack of asthma, which recurred when he again went into the park, and a skin test gave a positive reaction to horse dandruff diluted 1–20,000 but not to any of numerous other proteins. Also we have the passive transfer of sensitivity to milk when the blood of infants with exudative diathesis is injected into guinea pigs (Schlutz and Larson, 74).

But as soon as we leave such cases as these, and consider ordinary serum sickness with its cutaneous manifestations and other symptoms which follow at an uncertain period of time after a serum injection, often when this is the first injection of such a serum, the resemblance to anaphylaxis is much less distinct. Coca seems to have little hesitation in eliminating serum sickness entirely from the category of anaphylaxis reactions, laying particular emphasis upon the many points serum disease has in common with drug hypersensitiveness, which he believes cannot possibly be true anaphylaxis. He says "serum allergy exhibits no peculiarity that can serve as a criterion of separation of this condition from that of drug allergy and there is, thus, no occasion offered by the manifestation of serum allergy to seek an underlying mechanism that would be different from that operative in drug allergy, that is, one dependent on an antibody-antigen reaction." Hence he eliminates from the domain of anaphylaxis all the conditions manifested by cutane-

ous hypersensitivity, including the tuberculin reaction, hay fever, food and drug idiosyncracies, etc.

Kolmer (87), on the other hand, in his review on the cutaneous reactions, states that there are true anaphylactic skin reactions, "due to the interaction of a specific anaphylactic antibody and specific anaphylactogen largely within or without the cells, and with the formation of diffusible irritant capable of producing acute hyperemia, edema and leucocytic infiltration of the skin." He thus accepts the theory of local anaphylaxis depending on the formation of a poison by cellular reactions to foreign proteins introduced into the skin.

The relationship of the cutaneous manifestations of serum disease to anaphylaxis is especially well supported by the studies of C. W. Wells (88) on the precipitin content of the blood of patients who have received diphtheria antitoxin, for he found that in several persons who developed such manifestations the precipitin titer fell to rise again when the rash faded, as if the precipitin had been bound in the skin and thus caused a local anaphylactic reaction. Also, Weil (31) observed in human serum sickness a fall in blood pressure and a decrease in the coagulability of the blood (89), thus adding to the resemblance of this condition to true anaphylactic reactions.

However, among other facts that seem to separate cutaneous hypersensitiveness from systemic anaphylaxis, are the results described by Krause (90) in his studies on *The Nature of Resistance to Tuberculosis*. He states that whereas, as shown by Baldwin, it is easily possible to sensitize guinea pigs to tuberculo-protein so that they will give typical systemic anaphylactic reactions, such sensitized animals do not show cutaneous hypersensitiveness to tuberculo-protein, nor are they protected against infection with the tubercle bacillus. Skin sensitiveness can only be obtained by producing an actual tuberculous infection somewhere in the animal and when this tissue reactivity has developed it is found that the animal has the power to produce acute inflammatory reactions about injected tubercle bacilli, whereby its resistance is enhanced. Fleischner, Meyer and Shaw (91) have corroborated this and find that it is equally true for typhoid bacilli and B. abortus. Guinea pigs may be highly immunized with the dead bacilli or extracts therefrom, so that their blood has a high agglutinin content, or anaphylactic sensitization may be produced, and yet no cutaneous sensitivity can be shown; but guinea pigs infected with living bacteria give positive skin tests.

In support of the view that ordinary serum sickness following injection of foreign sera is not a true anaphylactic reaction, but is rather the result of a primary serum toxicity, may be cited the observation of Kraus (92) that bovine serum seldom causes serum sickness in man as compared with its frequency after horse serum. However, it is possible that man, using beef protein as a usual article of diet, may develop an actual immunity to beef proteins, as I found guinea pigs to do with the proteins forming the chief elements of their diet.

As to hay fever, Coca, who at one time described it as "a clinical expression of a local hypersensitiveness," the view most generally held, has been compelled by the evidence more recently developed, to take the opposite view and is now convinced that hay fever is not the result of a protein sensitization. He is satisfied to accept its hereditary character as an adequate explanation, although this does not in any way clear up the pathogenesis of the condition. The many observations of successful desensitization by repeated injection of pollen proteins is an outstanding fact which surely establishes a relation to anaphylactic sensitization, as also the close relationship to the many forms of asthma that are accompanied by skin hypersensitivity to the same proteins that are responsible for the asthma, and in which desensitization may be obtained by injection of the specific exciting protein. On the other hand, desensitization is accomplished by no means so readily or so surely as in typical anaphylaxis, and even marked systemic reactions produced by relatively large desensitizing doses of pollen may fail to desensitize (Rackemann, 93). An additional point, of which much is made in discussing the relation of hay fever to true anaphylaxis, is that it is usually impossible to confer passive hypersensitivity for pollen to guinea pigs by injecting them with the serum of patients with hay fever, but this negative fact seems of little significance if we consider that even the serum of guinea pigs highly sensitized to a foreign protein usually contains too few free antibodies to confer passive sensitization to other guinea pigs. The cells of a man may well be highly sensitized to a foreign protein without there being free antibodies in his blood to produce passive sensitization.

A new feature of local reactions has recently been brought to light by Auer (97), which may throw light on some of the manifestations of cutaneous hypersensitivity. Observing that operation wounds made in sensitized dogs at the time of reinjection showed a marked inflammatory reaction with a brawny edema, the suggestion was followed up, and the following facts were observed. If the ear of a normal rabbit

is rubbed gently with xylol there occurs a moderate inflammatory reaction which disappears in a short time without serious injury to the skin. The effects are quite the same in a rabbit that has been sensitized to a foreign protein, but in rabbits that have been sensitized and then reinjected with the same protein the xylol commonly produces a violent reaction, with exfoliative dermatitis followed by dry gangrene of the tips of the ears. The explanation of this striking effect seems to be simple. The slight inflammation produced by the xylol leads to a certain amount of inflammatory exudate. In the sensitized animals which have been recently re-injected with the sensitizing protein the blood contains more or less free antigen, which is poured out into the tissues with the exudate. Here it produces a local reaction with the sensitized tissues quite the same as if it had been locally injected, resulting in the usual local reaction seen in sensitized animals. Presumably similar effects could occur in any other tissue or organ. The importance of this observation lies in the recognition of a hitherto unappreciated mechanism by which anaphylactic reactions may be caused under certain conditions, by substances which themselves do not produce anaphylaxis. Conceivably, sensitized persons at times have free antigen in their blood, especially in cases of food sensitization, and if this is so, injury to any tissue or organ might possibly determine local anaphylactic reactions in these places.

BIBLIOGRAPHY

(1) von Pirquet, C.: Arch. Int. Med., 1911, vii, 259.
Hektoen, L.: Journ. Amer. Med. Assoc., 1912, lviii, 1081.
Besredka, A.: Anaphylaxis and antianaphylaxis and their experimental foundations; English ed., by S. R. Gloyne, 1919, St. Louis.
Doerr, R.: Kolle und Wassermann's Handb., 1913, ii, 947.
Zinsser, H.: Arch. Int. Med., 1915, xvi, 223.
Coca, A. F.: Tice's Practice of medicine, Prior & Co., New York, 1920, 107.
Boughton, T. H.: Journ. Lab. Clin. Med., 1920, v, 597.
Wells, H. G.: Chemical pathology, 4th ed., Philadelphia, 1920, 191.
(2) Karsner, H. T., and Hanzlik, P. J.: Journ. Pharm. Exper. Therap., 1919, xiv, 229, 379, 425, 449, 463, 479.
(3) Gay, F. P., and Robertson, T. B.: Journ. Exper. Med., 1913, xviii, 535.
Schmidt, C. L. A.: Univ. Cal. Publ. Pathol., 1916, ii, 157.
(4) Wells, H. G.: Zeitschr. f. Immunität., 1913, xix, 599. ("Betanucleoproteins" seem to have a thermostable antigenic capacity; see Journ. Biol. Chem., 1916, xxviii, 11.)
(5) Schmidt, C. L. A., and Bennett, C. B.: Journ. Infec. Dis., 1919, xxv, 207.
(6) Starin, W. A.: Journ. Infec. Dis., 1918, xxiii, 139; corroborated by Kahn and McNeil: Journ. Immunol., 1918, iii, 277.

(7) Wells, H. G., and Osborne, T. B.: Journ. Infec. Dis., 1911, viii, 66.
(8) Wells, H. G.: Journ. Infec. Dis., 1908, v, 449; Journ. Amer. Med. Assoc., 1908, l, 527.
(9) Wells, H. G.: Journ. Infec. Dis., 1909, vi, 506; 1911, ix, 147.
(10) Wells, H. G., and Osborne, T. B.: Journ. Infec. Dis., 1915, xvii, 259.
(11) Dochez, A. R., and Avery, O. T.: Journ. Exper. Med., 1917, xxvi, 477.
(12) Elliott, C. H.: Journ. Infec. Dis., 1914, xv, 501.
(13) Schmidt, C. L. A.: Journ. Immunol., 1920, v, 258.
(14) Ten Broeck, C.: Journ. Biol. Chem., 1914, xvii, 369.
(15) Landsteiner, K.: Biochem. Zeitschr., 1919, xciii, 106.
(16) Fink, E. B.: Journ. Infec. Dis., 1919, xxv, 97.
(17) Abderhalden, E.: Zeitschr. physiol. Chem., 1912, lxxxi, 315.
(18) Zunz, E.: Biochem. Journ., 1916, x, 160; Journ. physiol. et path. gén., 1917, xvii, 449; Arch. internat. de physiol., 1919, xv, 179, 192.
(19) White, B.: Journ. Med. Res., 1914, xxx, 393.
(20) Lake, G. C., Osborne, T. B., and Wells, H. G.: Journ. Infec. Dis., 1914, xiv, 364.
(21) Weil, R.: Journ. Immunol., 1916, i, 19.
(22) Coca, A. F., and Kosakai, M.: Journ. Immunol., 1920, v, 297.
(23) Zinsser, H.: Journ. Exper. Med., 1913, xviii, 219.
(24) Larson, W. P., and Bell, E. T.: Journ. Infec. Dis., 1919, xxiv, 185; see also Coca, A. F.: Journ. Immunol., 1919, iv, 209.
(25) Weil, R.: Journ. Med. Res., 1913, xxx, 87.
(26) Coca, A. F.: Zeitschr. f. Immunität., 1914, xx, 622.
(27) von Fenyvessy, B., and Freund, J.: Zeitschr. f. Immunität., 1914, xxii, 59.
(28) Weil, R.: Journ. Med. Res., 1913, xxx, 299.
(29) Manwaring, W. H., and Kusama, Y.: Journ. Immunol., 1917, ii, 157.
(30) Bronfenbrenner, J.: Journ. Exper. Med., 1915, xxi, 480.
(31) Weil, R.: Journ. Immunol., 1917, ii, 399.
(32) Dale, H. H.: Journ. Pharm. Exper. Therap., 1913, iv, 167.
(33) Abel, J. J., and Kubota, S.: Journ. Pharm. Exper. Therap., 1919, xiii, 243.
(34) Dale, H. H.: Johns Hopkins Hosp. Bull., 1920, xxxi, 257, 310.
(35) Smith, M. I.: Journ. Immunol., 1920, v, 239.
(36) Elsesser, O. J.: Journ. Infec. Dis., 1916, xix, 655.
(37) Cannon, W. B.: Compt. Rend. Soc. Biol., 1918, lxxxi, 850.
Delbet: Rev. de Chir., 1919, xxxviii, 309.
(38) Novy, F. G., and De Kruif, P. H.: Journ. Amer. Med. Assoc., 1917, lxviii, 1524; this being a summary of ten articles in the Journ. Infec. Dis., 1917, xx, 499 to 854.
(39) Sachs, H.: Kolloid Zeitschr., 1919, xxiv, 113.
(40) Danyz, J.: Journ. Infec. Dis., 1918, xx, 427.
(41) Coca, A. F.: Journ. Immunol., 1919, iv, 219.
(42) Weil, R.: Journ. Immunol., 1917, ii, 429.
(43) Rumpf, F.: Zeitschr. Immunität., 1918, xxvii, 489.
(44) Simonds, J. P.: Journ. Amer. Med. Assoc., 1919, lxxiii, 1437.
(45) Auer, J., and Robinson, G. C.: Journ. Exper. Med., 1913, xviii, 450.
(46) Pelz, M. D., and Jackson, D. E.: Journ. Lab. Clin. Med., 1918, iii, 387.
(47) Werigo, B.: Compt. Rend. Soc. Biol., 1916, lxxiv, 87.

(48) CORPER, H. J.: Journ. Infec. Dis., 1919, xxv, 248.
(49) JEANBRAU: quoted by GIRAUD: Arch. Méd. et Pharm. militaire, 1918, lxxix, 369.
(50) BÖTTNER, A.: Deutsch. Arch. klin. Med., 1918, cxxv, 1.
(51) HUBER, H. L., AND KOESSLER, K. K.: Arch. Int. Med., 1921.
(52) GURD, F. B., AND EMRYS-ROBERTS, E.: Lancet, 1920, cxcviii, 762.
(53) WELLS, H. G.: Centralbl. allg. Pathol., 1912, xxiii, 945.
(54) LONGCOPE, W. T.: Journ. Exper. Med., 1913, xviii, 678; 1915, xxii, 793.
(55) BOUGHTON, T. H.: Journ. Immunol., 1919, iv, 213.
(56) BELL, E. T., AND HARTZELL, T. B.: Journ. Infec. Dis., 1919, xxiv, 618.
(57) WEIL, R.: Journ. Immunol., 1917, ii, 525.
(58) BULGER, H. A.: Journ. Infec. Dis., 1918, xxiii, 522.
(59) MANWARING, W. H., AND CROWE, H. E.: Journ. Immunol., 1917, ii, 517.
(60) FALLS, F. H.: Journ. Infec. Dis., 1918, xxii, 83.
(61) LOEB, L.: Journ. Immunol., 1917, ii, 557.
(62) MAJOR, R. H.: Deutsch. Arch. klin. Med., 1914, cxvi, 248.
(63) AUER, J., AND VAN SLYKE, D. D.: Journ. Exper. Med., 1913, xviii, 210.
(64) BARGER, G., AND DALE, H. H.: Biochem. Journ., 1914, viii, 670.
(65) HISANOBU, K.: Amer. Journ. Physiol., 1920, l, 357.
(66) MANWARING, W. H., KUSAMA, Y., AND CROWE, H. S.: Journ. Immunol., 1917, ii, 511.
(67) WEIL, R.: Journ. Immunol., 1917, ii, 469.
(68) WEIL, R.: Journ. Med. Res., 1913, xxx, 299.
(69) WEIL, R.: Zeitschr. Immunität., 1913, xx, 199.
(70) THOMSEN, O.: Zeitschr. Immunität., 1917, xxvi, 213.
(71) LEWIS, J. H.: Journ. Infec. Dis., 1915, xvii, 241.
(72) WELLS, H. G.: Journ. Infec. Dis., 1911, ix, 147.
(73) BESREDKA, A.: Ann. Inst. Pasteur, 1909, xxiii, 166.
(74) SCHLUTZ, F. W., AND LARSON, W. P.: Arch. Pediatrics, 1918, xxxv, 705.
(75) KOPACZEVSKI, W., AND VAHRAM, A.: Compt. Rend. Acad. Sci., 1919, clxix, 250.
(76) RICHET, C., BRODIN, P., AND ST. GIRONS, F.: Compt. Rend. Acad. Sci., 1919, clxix, 9.
(77) OSBORNE, T. B., AND WAKEMAN, A. J.: Journ. Biol. Chem., 1918, xxxiii, 7 and 243.
(78) DALE, H. H., AND HARTLEY, P.: Biochem. Journ., 1916, x, 408.
(79) KATO, Y.: Mitt. Med. Fak. Univ. Tokio, 1917, xviii, 195.
(80) WELLS, H. G., AND OSBORNE, T. B.: Journ. Infec. Dis., 1916, xix, 183.
(81) WELLS, H. G., AND OSBORNE, T. B.: Journ. Infec. Dis., 1913, xii, 341.
(82) LANDSTEINER, K., AND LAMPL, H.: Biochem. Zeitschr., 1918, lxxxvi, 343.
(83) OBERMAYER, FR., AND PICK, E. P.: Wien. klin. Wochenschr., 1906, xix, 327.
(84) DAKIN, H. D., AND DALE, H. H.: Biochem. Journ., 1919, xiii, 248.
(85) KOSSEL, A.: Zeitschr. physiol. Chem., 1913, lxxxviii, 163.
(86) RAMIREZ, M. A.: Journ. Amer. Med. Assoc., 1919, lxxiii, 984.
(87) KOLMER, J. A.: Johns Hopkins Hosp. Bull., 1917, xxviii, 163.
(88) WELLS, C. W.: Journ. Infec. Dis., 1915, xvi, 63.
(89) SHATTUCK, H. F.: Arch. Int. Med., 1917, xx, 167.
(90) KRAUSE, A. K.: Amer. Rev. Tuberc., 1917, i, 65.

(91) KLEISCHNER, E. C., MEYER, K. F., AND SHAW, E. B.: Amer. Journ. Dis. Chil., 1919, xviii, 577.
(92) KRAUS, R.: Rev. del Inst. Bact. Buenos Aires, 1919, ii, 1.
(93) RACKEMANN, F. M.: Boston Med. Surg. Journ., 1920, clxxxii, 295.
(94) KRITCHEWSKY, J. L.: Journ. Infec. Dis., 1918, xxii, 101.
(95) HASKINOTO, M., AND PICK, E. P.: Arch. exper. Path. u. Pharm., 1914, lxxvi, 89.
FENYVESSY, B., AND FREUND, J.: Biochem. Zeitschr., 1919, xcvi, 223.
(96) HANKE, M. T., AND KOESSLER, K. K.: Journ. Biol. Chem., 1920, xliii, 521–579.
(97) AUER, J.: Journ. Exper. Med., 1920, xxxii, 427.

PHOTO-ELECTRIC CURRENTS IN THE EYE

CHARLES SHEARD

Southbridge, Massachusetts

The effect of the stimulus of light on the retina is perceived by the brain as a visual sensation. The process or processes by which the ether-wave disturbance causes this visual impulse is still very obscure. As a matter of fact the whole of the field of photo-chemical action is still in its infancy. The process of making an ordinary negative by exposing a dry plate in a camera to light and the manner of developing and fixing such a plate are mechanically easy of accomplishment. But ever since the discovery that an invisible light effect could be developed into a strong image by the application of suitable reducing agents, the constitution of the invisible or so-called latent image has been the subject of study and controversy and no wholly satisfactory explanation of the effects of radiation upon silver salts has been presented.

The process by which the ether disturbance causes a visual impulse may be ascribed to *a*, chemical action, *b*, molecular strain and *c*, electrical action.

According to the *chemical* theory it is presumed that certain visual substances (or possibly substance) in the retina are affected by light and that vision originates from metabolic changes produced in these visual substances. It is supposed that the metabolic changes consist of two phases; the upward, constructive or anabolic phase, and the downward, destructive or katabolic change. These anabolic and katabolic changes in various visual substances are supposed to produce the variations of sensation of light and color. This theory is complex. Numerous objections have been urged against its acceptance; for it is difficult, for instance, to see how this very rapid visual process can be due to a comparatively slow chemical action consisting of the destructive breaking-down of the substance followed by its renovation. Support was at first furnished the chemical theory as it has ordinarily been understood and presented by the bleaching action of light on the visual purple present in the retina, but it has been discovered that the presence or absence of visual purple is not essential to vision and that its presence is of only secondary importance. For it is well known that the

visual purple is lacking in the fovea centralis and it is also found to be completely absent from the retinae of many animals possessing keen sight. Edridge-Green (1) states his belief as to the nature of retinal stimulation in the following words:

> A ray of light impinging upon the retina liberates the visual purple from the rods and a photograph is formed. The rods are concerned only with the conveyance of the light impulses to the brain. The ends of the cones are stimulated through the photo-chemical decomposition of the visual purple by light, and a visual impulse is set up which is conveyed through the optic nerve fibres to the brain. The character of the stimulus and impulse differs according to the wave-length of the light causing it. In the impulse itself we have the physiological basis of the sensation of light and in the quality of the impulse the physiological basis of the sensation of color.

The *mechanical* theory depends, in large measure, upon the theory of resonance in connection with chemical action. It is readily conceivable that a ray of light can cause a chemical decomposition of a substance in which the rhythmic excursions of an atom or atoms from, or round about, the center of attraction in a molecule are in exact tune with the waves of light falling on such atoms. The excursions may be so increased in extent by the rhythmic energy supplied by the light waves that the atoms will leave the parent molecules and produce new molecules. It is not as easy to see why the rhythmic excursions of atoms in the same molecule are also increased to the point of molecular rupture when the wave-motion of the impinging rays is not in tune or nearly so. But some photographic and mechanical examples help us out. For if a sensitive salt, such as silver chloride, is exposed to the action of the spectrum, we can plot a curve showing the sensitiveness of this particular salt to the different spectral rays. Such a plotted curve shows a rise in sensitiveness to a maximum followed by a decline; the maximum of such curves shows the place in the spectrum where the vibrations composing the beam of light are in tune with the vibrations of the chlorine atom in silver chloride, for example, the chlorine being that part of the molecule which is swung away and annexed to some other adjacent molecule. We also have mechanical examples of the effects produced by vibrations which are not in tune with, but which act upon, a vibrating body. A simple apparatus consists of two different pendulums which can act upon one another through a proper communicating medium; such would occur, for example, when the pendulums are connected to a taut piece of rubber tubing fastened horizontally. When the pendulums are of equal length and one is started

into motion, the second one also begins to swing and, since it is in tune, the amplitude constantly increases. But if one pendulum is a little longer or shorter than the other, experimentation shows that one pendulum causes the second one to swing with increasing amplitude and by degrees the two will swing in opposite directions; the amplitude of the first pendulum will decrease and finally come to rest, when the motion starts over again as at first. Thus, if the mechanical analogy is applicable here, it is seen that if the waves causing a ray of light are out of tune with the vibrations of the atoms the amplitude will still be increased and the increase can be such as to swing the atom beyond the sphere of molecular attraction and so decompose the molecule but with less ease than when the waves are in tune. Houstoun (2) has lately attempted another explanation of color theory on the basis of the visibility curve of the eye. He believes that the visibility curve is explicable on the supposition that there exists in the eye a very large number of vibrators with a free period in the green and that these execute forced vibrations under the influence of the light wave, and that the amplitude of the forced vibration is a maximum when the free period of the vibrators coincides with the period of the incident light. When the energy of the vibrator reaches a critical value, the force attaching the vibrator to its center snaps, the latter then ceases to absorb light energy and a chemical change takes place.

We have written of the *resonance* theory as a correlation and interaction between radiant energy and the atom of the molecule and have thus followed and outlined the theory as it is generally presented. However, in the light of modern physics, we should presumably have written of resonance and electrons. In fact, the essential points in the mechanical theory of retinal stimulation as consisting of resonance effects coupled with chemical action fit in with many of the physical phenomena known as *photo-electric* actions. This term includes phenomena due to the action of light in liberating negative electrons from various metallic substances. It is known, for example, that there is a considerable influence of the wave-length or frequency of the light upon the number of electrons omitted and that curves plotted between frequency and rates of "leak" of negative electricity from metals such as sodium, potassium and rubidium show maximum or resonance effects. Likewise, salts which undergo decomposition in the light, such as silver chloride, are strongly photo-electric. We are, therefore, presumably dealing under the tenets presented to us in this theory with the expulsion of electrons due to resonance; the electrons are set in

resonant vibration by the incident light and acquire sufficient velocity to enable them to escape from the atom. In the case of photo-electric effects from metals it has been possible to identify the valency electrons which play the part of chemical bonds with the photo-electric electrons which can be liberated by the action of light. From this point of view a photo-electric change and a photo-chemical change may be regarded as of the same character, consisting essentially in the displacement or separation of an electron through the absorption of energy from a light wave. We may thus have two kinds of photo-chemical action. In a primary or direct action the same valency electron which has absorbed the energy is itself released from its connection with a strange atom, whilst in a secondary or indirect action the electron which has absorbed the energy causes by collision the separation of an electron from an atom. We may interpret some of the results obtained by Gotch (3), (4) in the light of these facts. From the viewpoint of modern science we may regard it as fairly certain that the first stage in any photo-chemical reaction consists in the separation of negative electrons under the influence of light. It is obvious that absorption of light is necessary before any change can be brought about. According to the theory of Grotthus (5) we may say that "the action of a ray of light is analogous to that of a voltaic cell:" that is, the chemical action of light is essentially electrolytic in nature.

These various fundamental concepts have been presented for the reason that there are various data afforded by the experimental results to be discussed further on in this paper which are in accord with certain of the principles underlying both the chemical and mechanical theories.

Hence the *electrical* theory supposes that the visual impulse is the concomitant of an electrical impulse; that an electrical current is generated in the retina under the influence of light and that this is transmitted to the brain through the optic nerve. It is an undoubted fact that light gives rise to retinal currents and that, on the other hand, an electrical current suitably applied causes the sensation of light. Holmgren, Dewar, M'Kendrick, Kühne, Steiner, Waller and others have shown that illumination produces electric variation in a freshly excised eye. The currents are very small in value, hence a very sensitive and suitable galvanometer must be used. Currents of injury or contact differences of potential arise when the galvanometer terminals are connected to the cornea and the cut end of the nerve respectively; these may be compensated by means of a potentiometer device. Normally

the current of response is from the nerve to the retina: hence these currents have been designated as positive when flowing from the retina to the cornea.

GENERAL EXPERIMENTAL PROCEDURE

The general experimental method adopted in investigations on the retinal currents due to light stimuli is illustrated in figures 1 (a) and 1 (b). Contacts are made with the galvanometer, *G*, through non-polarizing electrodes with the cornea of the enucleated eye on the one

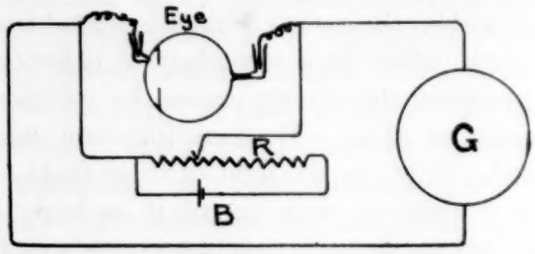

Fig. 1 (a). Detailed diagram of compensating potential device

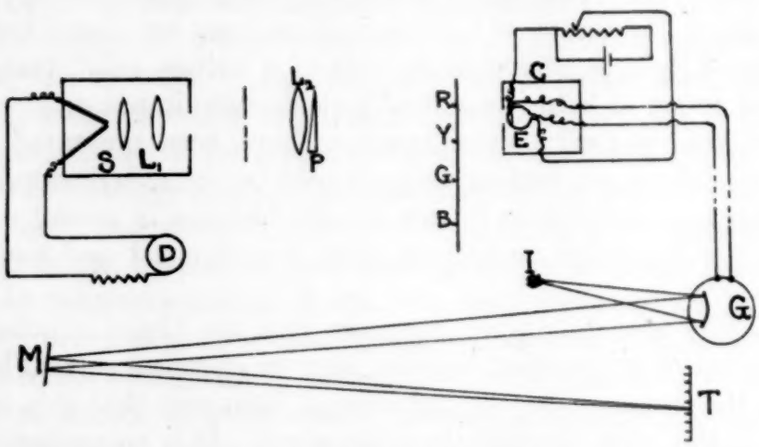

Fig. 1 (b). A satisfactory ensemble for experimental work

side and at the cut optic nerve on the other end. When an excised eyeball is connected up through non-polarizable electrodes in the manner indicated, a current of injury is found which must be compensated for: one method is diagrammed in figure 1 (a). This is not, however, the best method experimentally in that the eyeball only is included rather than having both eyeball and galvanometer in the compensating system. A constant voltage cell may be short-circuited

through a high resistance, R, from which any desired potential difference may be tapped off and applied in such a direction (by proper connection of the terminal points of the battery to the eyeball) as to bring the galvanometer reading back to its initial zero point, thus compensating for the effect known as the current of injury. In this manner, provided the compensating electromotive force demanded does not vary with time, the effects obtained through illumination are genuine and are not complicated by injury-effect currents or similar phenomena.

The schematic arrangement of apparatus used in 1918 by Sheard and McPeek (6) is shown in figure 1 (b). The source of light, S, was a carbon arc, hand feed, supplied by an alternating current dynamo, D. The quartz lens system, L_1, threw the light upon a slit and thence to a lens, L_2, and a prism, P, carrying on one face a grating replica. A continuous spectrum was formed upon a cardboard screen marked into wave-lengths and represented by R, Y, G, B, in figure 1 (b). The spectrum projection apparatus, which was fastened to a single supporting base, was turned so as to allow the eye to be exposed to any desired portion of the spectrum.

All eyes used should be freshly excised, otherwise they will give responses characteristic of what Waller (7) calls transitional changes. Likewise these eyes must not be kept for any length of time in Ringer's or similar solution, since they lose their sensitiveness of response. Great care is needed to prevent mechanical pressure or other injury to the eye to be used, otherwise responses which are not characteristic of the normal eye will exist.

Various types of galvanometer may be used. They must be extremely sensitive, however. Galvanometers of the d'Arsonval or Thomson type should have sensibilities of the order of 10^{-9} to 10^{-10} ampere. Gotch (3), (4) used a capillary electrometer which recorded the rapid alternations of current present in the responses of the eye. By this means the responses obtained and effects produced in small fractions of a second may be investigated; hence such a device is particularly suitable for investigating the lag of response behind luminous stimulus. Einthoven and Jolly (8) made use of a string galvanometer, which serves essentially the same purpose as the capillary electrometer.

GENERAL SURVEY OF EXPERIMENTATION AND RESEARCH

Holmgren (9) is accredited with the discovery that when light is allowed to fall upon the eye of a frog which has been kept in the dark and again when the light is removed, there is an increase in the posi-

tive direction of the current present during darkness. He found that the strength of these retinal currents is, within limits, proportional to the intensity of the light; also, that the onset and removal of light is attended with the same electrical changes when the posterior half of the eyeball only is illuminated.

Independently of Holmgren and at about the same time, Dewar and M'Kendrick (10) made investigations upon the electrical changes caused in retinae by light. The conclusions to which they arrived are: *a*, That the impact of light on the eyes of members of the following groups of animals, viz., mammalia, aves, reptilia, amphibia, pisces and crustacea, produces a variation amounting to from 3 to 10 per cent of the normal electromotive force existing between the corneal surface and the transverse section of the nerve; *b*, that this electrical alteration may be traced to the brain; *c*, that those rays which we regard as most luminous produce the largest variation; *d*, that the alteration of the electrical effects with varying luminous intensity seems to follow very closely ratios given by the psycho-physical law of Fechner; *e*, that the electrical alteration is due to the action of light on the retinal structure itself, as it is independent of the anterior portion of the eye, eliminating, therefore, the natural supposition that the contraction of the iris might produce such a result; and *f*, that it is possible by experiment to discover the physical expression of what is usually called in physiological terminology, fatigue.

Kühne and Steiner (11) investigated chiefly the reactions of the isolated retina. They found that the electrical change on lighting and darkening is a complex one. The reaction is divisible, they say, into two parts; the first one, due to the onset and continuance of illumination, consists of a negative variation preceded by a positive; and the second, caused by the disappearance of light, consists of a simple positive variation. No current was obtained from the posterior half of the eyeball from which the retina had been removed, the pigment layer being left behind. Hence the pigment epithelium possesses no electromotive force or power and the source of the changes is in the retina.

Waller (7), whose brochure on *The Signs of Life* is most fascinating, carried out a number of important researches in this field of investigation. A series of photograms of responses to light in the frog's retina is given in figure 2. These maxima represent five normal responses due to a candle at 2 feet with illumination for one minute and obscurity for two minutes respectively. The abscissae represent the time in minutes and the ordinates the absolute electromotive forces. These

results indicate rapid rises in electromotive force under illumination together with less rapid recovery.

The main conclusions to which Waller (12) arrived in his work are: *a*, A fresh normal eyeball manifests a positive current which gradually declines to zero and becomes reversed. By *positive current* Waller refers to the current of response due to the action of light on the retina which is from the nerve—that which is not directly stimulated by light—to the retina; that is, from the less excited to the more excited. *b*, On exposure to light the normal current, whether positive or negative, undergoes a positive variation. *c*, The magnitude of the response to

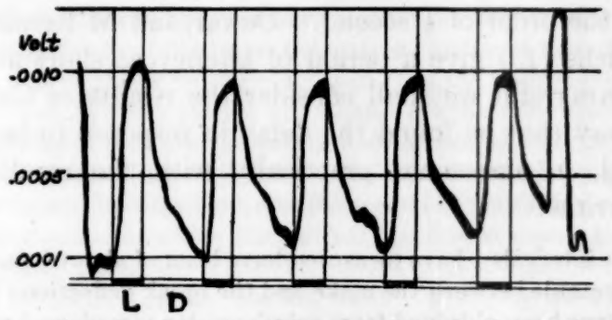

Fig. 2. A series of responses with a frog's eyeball. (After Waller)

light increases with the duration of illumination. *d*, The magnitude of the response to light increases with the strength of illumination. *e*, Colored lights act in the same direction and in accordance with their luminosity.

TYPICAL RESPONSES. TRANSIENT RISES OF CURRENT. LAG OF RESPONSE

A considerable portion of that which is to be presented in the next few paragraphs is due primarily to Waller. The conclusions are rather important and some of them far-reaching.

Typical responses. Figure 2 shows that, during the exposure of the eye to light, there is a current flowing from the fundus toward the cornea and, in the case illustrated, there is developed a potential of the order of $\frac{1}{1000}$ volt. On cessation of the light the response falls off asymptomatically to the value of about $\frac{1}{10000}$ volt. It will be noticed that there is evidently not a full recovery since there is a constantly rising minimum value at the end of each 2-minute period of darkness.

Transient rises of current. Another phenomenon observed by Kühne and Steiner and verified by Waller is that, immediately on the stoppage

of light, there is often a sudden increase in retinal current before the usual recovery takes place. These are spoken of as transient positive augmentation effects. An inspection of the results obtained shows that, on illumination, the response-curve rises, that continued illumination produces a decline, and that on the cessation of light there is a transient rise of current.

Lag of response. Furthermore, Waller found a considerable interval of time between the incidence of light and the electrical response; sometimes the delay at "make" of light might be so great as to be measurable with a stop-watch. Several records given by Waller show that there is a retinal delay at both the beginning and at the end of illumination of the order of $\frac{1}{6}$ second. Dewar and M'Kendrick (10) and still later Fuchs (13) give a period of latency as short as $\frac{1}{100}$ second. In a later paragraph we shall consider the results of Gotch (3), (4): suffice it to say that he found the delay in response to be of the order of $\frac{2}{100}$ second, in agreement practically with the results of Waller. Waller (7) writes:

> The shortest intervals I have measured have been of about $\frac{1}{3}$ second, no difference being detectable between the make and the break deflections in this respect. These values have been obtained from galvanometer records and are subject to a correction by reason of the physical lost-time of the suspended magnets and mirrors. With the capillary electrometer, which has no appreciable lost-time, the latency comes out at about 0.15 second—still a very considerable delay and indicative of a rather surprising physiological inertia of the retinal organ.

Three types of response. Three types of response to light are differentiated by Waller (7).

I. Positive response of the first type, characteristic of the normal fresh uninjured eyeball.

II. Mixed responses, characteristic of transitional states between I and II.

III. Negative responses of the third type, characteristic of the compressed or partially injured eyeball.

In the first or normal type, the response is a positive one, proceeding quite rapidly at first to a final "saturation" value, similar to the results shown in figure 2. In class II, on the stimulation by light there is an initial positive response, followed by a rapid falling off to decided negative potential values, which is in turn followed by positive currents on the cessation of light. In the reversed or third type, the stimulus of light gives an immediate negative electromotive effect. Waller discovered that the principal factor in bringing about that state of eyeball

in which the response to light is purely negative, is a moderate compression or shrinking by drying of its retinal coats. Gentle massage of the eyeball will cause a gradual conversion from type I to type II.

Source and character of the electromotive forces. Furthermore, all the experimental facts obtained, viz., the transitional responses from positive to negative, the immediate conversion from positive to negative by massage, the initial period of hesitation and the latency of response at "break," point to the conclusion that the retina when exposed to light is the seat of two contrary electromotive changes. In commenting on these effects Waller (7) writes:

> I think that anyone who has witnessed a great number of experiments of this character and who has carefully inspected a large number of records such as those figured above, will be forced to the conclusion that the normal positive response is the algebraic sum of two opposite and almost synchronous electromotive phenomena, one giving a positive current that over-compensates, the other giving negative current; the first more labile than the second, so that by gentle compression it may, so to say, be wiped out and the second be thus unmasked. A careful scrutiny of the manner in which the normal positive response begins and ends confirms this view, and at the same time offers to us a plausible explanation of what has been referred to as the "parting bow." If during the exposure to light there is a tug-of-war between positive and negative current, with predominance of the former, and if at the end of exposure both currents should cease, but the negative cease more rapidly than the positive, then we should witness what actually does happen, viz., a short movement in the positive direction preceding the return to a state of rest. And finally, on turning back to the nicer examination of the rising effect at the outset of exposure, we find another sign of an opposition between two contrary and all but synchronously developing currents. There is often at this point a false latent period, or period of hesitation, perceptible on simple observation or, better still, by means of records where the beginning of an exposure has been mechanically signalled, amounting to several seconds, and intelligible only on the supposition that our galvanometric magnet is, so to say, trembling in the balance between two opposite and almost perfectly congruent forces. . . . All these facts—viz., the transitional responses from positive to negative, the immediate conversion from positive to negative by compression, the initial period of hesitation, the terminal "parting bow"—point to one and the same conclusion, viz., that the retina, when exposed to light, is the seat of two contrary electromotive changes. And it matters little whether you imagine two single processes or one double process behind the movements of the machinery.

Cause and effect. The cause/effect relationship between stimulation by light and retinal change has been investigated by Holmgren (9), Dewar and M'Kendrick (10) and Waller (12). Waller used a light of suitable strength, varied on an arithmetic scale constructed in con-

formity with the law that luminosity varies inversely as the distance squared. He found that the curve showing the relationship between the units of excitation as abscissae and the voltages of response as ordinates was concave to the abscissa. Such a relationship is in conformity with almost every series of observations we are able to carry out on living matter. Muscular contraction and the negative variation of a nerve-current give the same general type of curve. It is apparently a general law that the response of living matter to the excitation of its environment increases by diminishing increments, giving a cause/effect curve that is concave toward its abscissa.

Effects of tetanization. Waller (12) also investigated the effects and after-effects of tetanization upon the eyeball as well as the influence of tetanization upon its electrical responses to light. This gave rise to his work on so-called "blaze currents." Tetanizing currents were produced by a small induction coil, the secondary of which was in circuit with the galvanometer and eye. A marked increase in electrical response to light was found after tetanization. This marked augmentation of response as to current may be due to augmentation of voltage or of conductivity or to both factors. Waller believes that it is due to both factors, the electromotive force of the response being increased and the resistance of the eyeball being decreased. Marked effects not only follow after tetanization but a similar effect occurs during this procedure.

Anesthetics and retino-motor effects. Engelmann and Grijns (14) thoroughly investigated the retino-motor effects of anesthetics upon retinal currents. It was found that ordinary anesthetics—carbon dioxide, ether and chloroform—influence the electrical response of the retina to light as might be expected from a consideration of the physiological character and conditions of the response and the relative power of the anesthetic used. Using an enucleated eyeball, the layer to be anesthetized is comparatively well protected by the sclerotic coat, and the effects of an anesthetic vapor are obstructed if only because its access to the retina is obstructed. However, the characteristic effects of the three anesthetics are produced although more slowly than in the case of an isolated nerve. Carbon dioxide gives diminution followed by an augmentation of the response. Ether gives temporary diminution or abolition of the response, followed ordinarily by perfect recovery. Chloroform gives abolition of the response and when once this abolition occurs it is usually final.

These effects as studied by Engelmann and Grijns are of interest in two particulars, as pointed out by Waller (7). They indicate the pos-

sible existence of efferent fibers in that most typical afferent nerve, the optic nerve. Also they afford some answers as to whether these electrical changes belong to altered pigment or retracted cones or both phenomena. Engelmann's observations indicate that the retraction of the cones and an electrical effect can be produced in one eye by electrical excitation of the peripheral end of its optic nerve, or of the central end of the other optic nerve.

LATER EXPERIMENTATION

Gotch (3), (4) experimented with frogs' eyes and made use of a capillary electrometer which recorded the rapid alterations of current present in the retinal responses. In particular, he investigated the time relations of the photo-electric changes. Especially does he refer to the researches of Fuchs (13) in that the delay between illumination and eyeball effect is of the order of a few thousandths of a second. The method employed by Fuchs, although valuable for nerve and muscle, is singularly inappropriate for the investigation of retinal effects. Fuchs used a sparking rheotome, the period of time of which was such as to be of service with muscle and nerve but far too rapid for photo-electric effects.

Professor Gotch (3) concludes in his paper of 1903 that: 1, The curves show for the first time the character of the time relations. These comprise: *a*, a period of delay before the change due to illumination occurs, varying slightly with the temperature, but at 16°C. lasting on an average 0.17 to 0.18 second; *b*, a period of development lasting until from 0.6 to 0.8 second with persistent illumination, and reaching a maximum varying from 0.0004 volt to 0.001 volt; *c*, a period of subsidence to a certain sustained level; *d*, a second period of development when the light ceases commencing 0.15 to 0.17 second after the cessation and developing more rapidly than the initial rise due to the onset of light; *e*, a period of continued slow subsidence of the terminal effect during the darkness which succeeds the illumination. 2, The photo-electric effect of a flash of light develops more slowly and attains a much lower maximum than that produced by sustained illumination. 3, The effect of the cessation of light is largely determined by the duration of the previous exposure. When the duration is optimum for the particular eyeball employed, the effect produced by light cessation is far more pronounced than any change caused by illumination. 4, The favorable condition for obtaining an effect upon sudden illumination is

previous darkness. 5, The photo-electric changes are all of the same general type, giving rise to monophasic effects and appear to be mainly if not entirely due to processes occurring in the posterior part of the eyeball, presumably the retinal and pigmentary layers. 6, The results appear to indicate the localization of two photo-chemical substances in the posterior half of the eyeball, these being a substance reacting to light and a substance reacting to darkness. Each reaction is a change of the same type, but the substances must undergo phases of metabolism under conditions opposite to those which evoke the reaction effects.

In the second paper of 1904 Gotch (4) summarizes his work as follows: 1, These photo-electric responses fail or become extremely feeble if the eyeball is placed in the infra-red or ultra–violet regions of the spectrum. 2, The range of light vibrations within which the frog's eyeball gives definite photo-electric responses corresponds very closely to the range of vision in the case of our own color sensations. 3, All responses are of the same general type, whether they are evoked by white or by colored light. 4, An additional response is obtained when the light is suddenly replaced by darkness; this is of the same fundamental type as the illumination response. 5, There is no evidence of any excitatory process except those of the fundamental type. 6, The time-relations of the responses evoked by the various colored lights and by darkness are not identical. 7, There are four responses independent as regards their causation: *a*, the response to red light characterized by a long latency nearly $\frac{3}{10}$ second and by its attaining a considerable maximum averaging about 0.0004 volt; *b*, the response to green light characterized by the same short latency as that found in the response to white light; it is also characterized by its maximum averaging 0.0005 volt; *c*, the response to violet light characterized by a latency longer than that of the green response but distinctly shorter than that of the red; it has low intensity, the maximum reached averaging only 0.00024 volt; *d*, the response to sudden darkness, characterized by a remarkable constant latency of not more than $\frac{2}{10}$ second. 8, The results appear to be in accordance with the theory of Young-Helmholtz as modified by Maxwell; in addition they appear to indicate that the eye reacts to sudden darkness. 9, Furthermore, the latent period of the photo-electric reaction is in a high degree dependent upon the intensity of the stimulus. With strong stimuli it is of the order of 0.01 second, while with very weak stimuli it may be lengthened to more than 2 seconds. These values are in agreement with the latent periods of light perception in the human eye.

Other important researches, in a general way confirmatory of certain of the conclusions previously arrived at, have been made by Fuchs (13), deHaas (15), Piper (16), Ishihara (17), Brücke and Garten (18). The latest paper of which the writer has any knowledge (although the paper is not at hand in preparing this article) is by Reidel (19). He measured the photo-electrical changes in the eye of the lobster and found continued positive deviation in the light-adapted eye and a steep positive initial deviation with quick falling off in the dark-adapted eye. These effects are in accordance with what we should expect if the retinal currents are at all analogous to photo-electric currents from metal surfaces or the discharge of electrons from any surface. For the light-adapted eye would have reached the condition of slow change with time in which the processes of ionization and recombination would be comparable. In the case of the dark-adapted eye, however, a rapid decay of current with time would be expected, since the amount of substance or substances easily ionizable would be appreciable but would be used up very rapidly.

Sheard and McPeek (6) have investigated in some detail the character of the photo-electric reactions to light stimuli of different wave-lengths as well as to the forms of the retinal current-time responses produced by seriatim exposure of light stimuli of different wave-lengths.

The form of the photo-electric reactions to the initial light stimuli of different wave-lengths. A freshly excised dog eye was taken and, after preliminary arrangements, was initially exposed to a region of the spectrum such that λ5900 (sodium line) fell upon the center of the pupil. The total time of exposure was 4 minutes. The plotted results are shown in figure 3. The difference between the initial electromotive force (just before exposure, which was zero) and the final zero or reading after 4 minutes is +10 divisions or +0.0016 volt. The eye was then screened from the radiation but the process apparently continued such that in the end the final reading after 6 minutes of darkness was 30 divisions above the initial zero or +0.0048 volt. The lines drawn through the *A*, *B*, *C*, etc., and the minimum points *D*, *E*, *H*, etc., are parallel and show the gradual *rise*, from the fiducial or arbitrary balance point, of the potenial difference under exposure to the yellow (sodium light) region of the spectrum. This would indicate the constant growth of a positive electromotive force independent of the superposed maximum responses.

In contradistinction to these results obtained with λ5900, figure 4 shows the form of responses and potential differences when another

freshly excised eye was initially exposed for $4\frac{1}{2}$ minutes to the violet λ4310. The color represented by this wave-length is roughly complementary to that of λ5900, the effects of which have been described in conjunction with figure 3. The maximum value of the electrical response to λ4310 (at center of pupil) taken at any point during the first 2 or 3 minutes as judged with respect to its own minimum is practically −0.0024 volt. The difference between the initial electromotive force—which was zero under the compensating potential device—and the final reading after 4 minutes' exposure was practically −0.0024 volt. The eye was then screened from radiation and the end reading (after 6

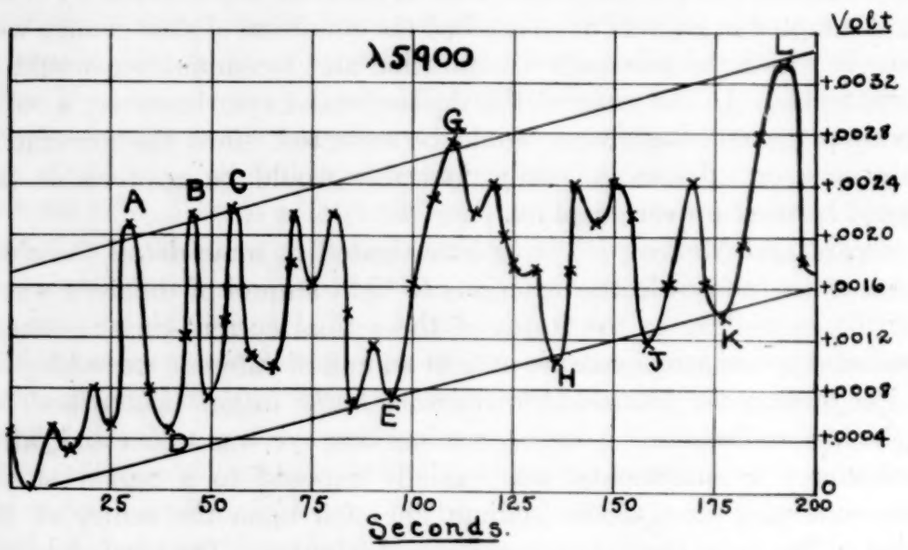

Fig. 3. Potential-time curve due to stimulation by λ5900

minutes' rest) obtained, indicating a total difference between the initial and the end reading (over the whole period of time) of −0.0019 volt. This is less than the general indicated difference, up to 3 minutes' exposure, of −0.0024 volt. However, the concluding portion (from abscissa 180 to 270 seconds) shows a tendency for the contrary or opposite electrical effect (?) to exert itself more pronouncedly. The lines drawn through the points *A*, *B*, *C* and the maximum, electrically considered, points *D*, *E* and *F* are parallel and show a gradual *fall* from the arbitrary balance starting point under exposure to λ4300. This would indicate the growth of a *negative* electromotive force independent of the superinduced fluctuations giving rise to maxima and minima.

In the case of exposure to λ5890 the fluctuations of potential difference are always in such a direction that the maxima indicate *increased positive potentials at the nerve*, with returns to minima which are still always positively above the fiducial zero. In the case of exposure to λ4310, however, the fluctuations of potential difference are always in a direction indicative of *increased positive potentials at the retina*, hence *reversed* potentials as compared with the first named set of conditions.

The forms of the retinal responses produced by seriatim exposure to light stimuli of different wave-lengths. Figures 3 and 4 give initial effects in two different eyes. In figure 5 have been plotted seriatim the results

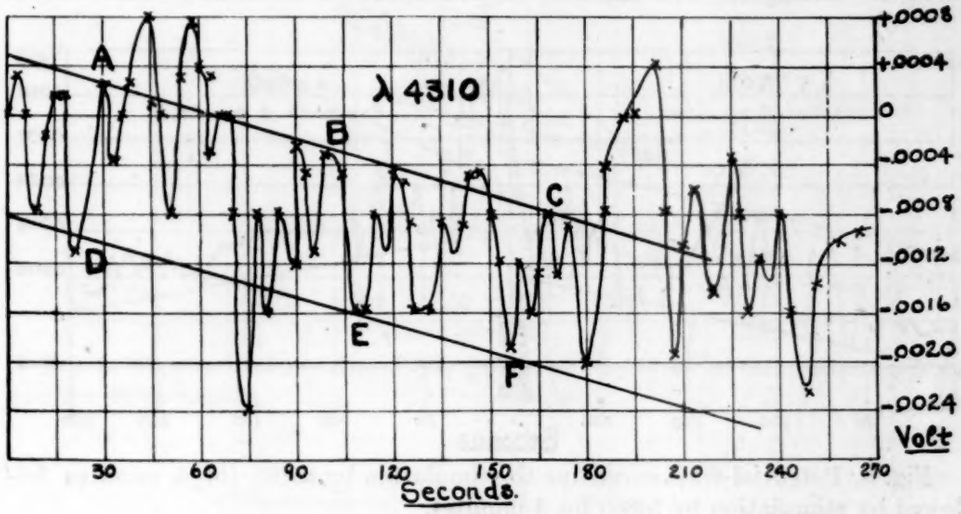

Fig. 4. Potential-time curve due to stimulation by λ4310

obtained by exposure of the same eye to λ5900, first, for a period of 4 minutes, a period of 6 minutes' rest, and a second exposure to the blue-violet λ4860, the complementary color to λ5900, for a period of 4 minutes. The same general conclusions are to be drawn from the seriatim exposures as from the individual exposures in separate eyes. In the case of exposure to λ5890 there is a maximum rise of electromotive force in a positive direction of 0.003 volt; the difference between maxima and minima in the fluctuations of potential is 0.0024 volt (practically); there is a steady rise in the permanent electromotive force during the continuance of the illumination amounting to about 0.0012 volt. After the 6 minutes' rest, the initial exposure to λ4860 produced but slight, although decreasing, potential effects during the first 25 seconds, to be immediately followed by an increased electromotive force in an oppos-

ing direction such as to carry back the positive potential established under illumination by λ5890 toward the initial zero. The conflict of the two processes is very clearly indicated in the record of the responses to λ4860. The curve as a whole shows the establishment of a contrary electromotive force as compared with that established by λ5890; a decided reaction to λ4860 at the end of 25 seconds' exposure and the gradual decrease of the positive potential difference established under stimulation by λ5890.

Figures 6 and 7 present in graphical form the experimental results obtained when an eye was exposed to luminous stimuli of various wave-lengths seriatim. The manner of procedure in obtaining these curves

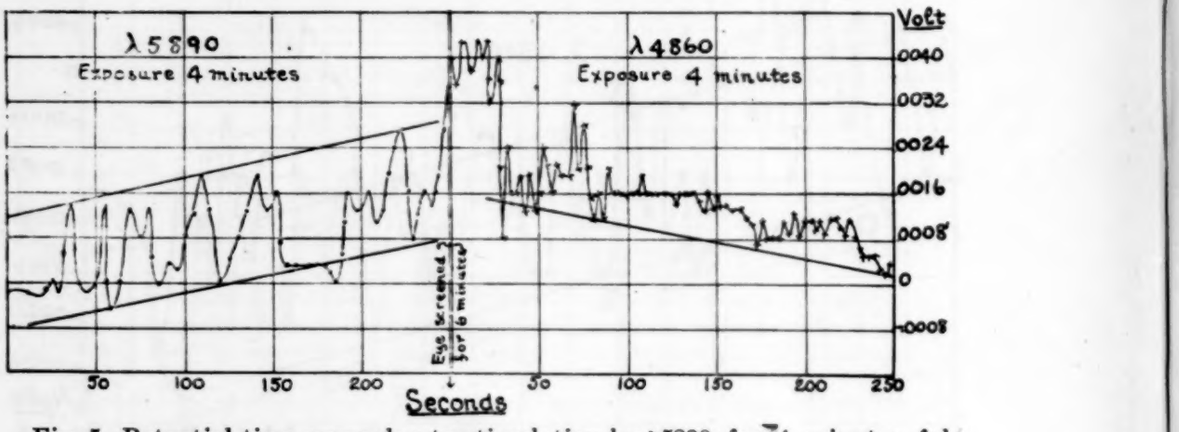

Fig. 5. Potential-time curve due to stimulation by λ5890 for 4 minutes followed by stimulation by λ4860 for 4 minutes.

was as follows: A fresh eye, an hour after excision, was connected into the galvanometer circuit and the galvanometer needle brought back to its original zero through the agency of the compensating potential device. The eye was then exposed to a certain region of spectral energy for 4 minutes after which the eye was again stimulated by luminous energy of entirely different wave-length. This process was continued over several such steps in the case of each eye investigated. The initial stimulation was obtained from entirely different spectral regions in the four curves presented in figures 6 and 7 These curves are plotted with the abscissae indicating the region of light stimulation: the ordinates represent the change in potential as registered by the galvanometer at the end of each cycle, namely the 4 minutes of exposure and the 6 minutes of rest. The order of stimulation, and regions used therefor are indicated by the serial numbers *1*, *2*, etc., or by letters *a*, *b*, *c*. In

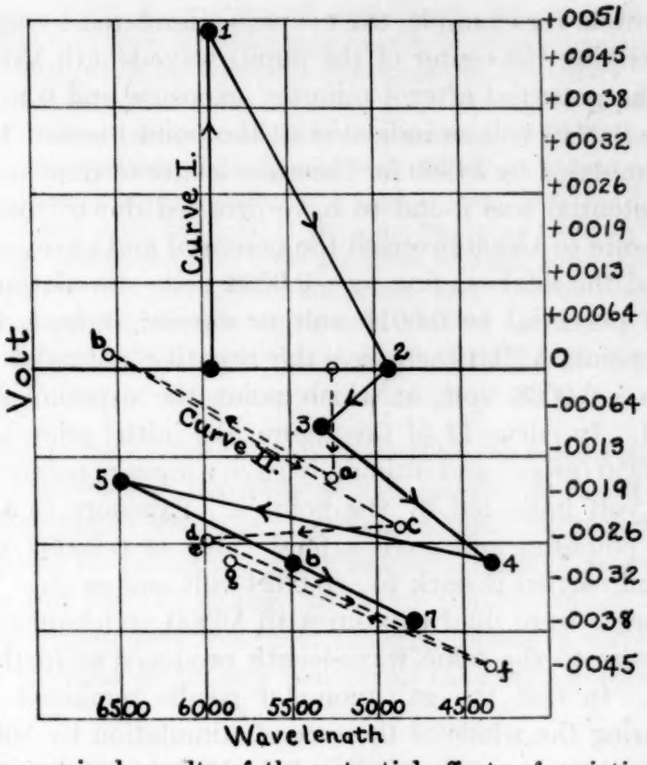

Fig. 6. Summarized results of the potential effects of seriatim exposures to various wave-lengths.

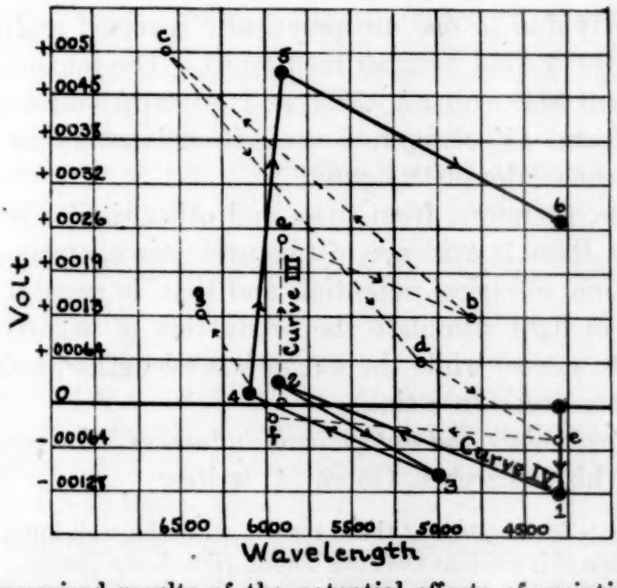

Fig. 7. Summarized results of the potential effects of seriatim exposures to various wave-lengths.

curve *I*, figure 6, for example, the eye was stimulated by light of average (i. e., striking the center of the pupil) wave-length λ5890 (sodium yellow). The potential after 4 minutes' exposure and 6 minutes' rest had risen to 0.0049 volt as indicated at the point marked *1*. The eye was then stimulated by λ4860 for the same length of time and occluded, when the potential was found to have dropped down from 0.0048 to zero. Exposure to λ5300 reversed the potential and gave −0.007 volt; λ4350 carried the final reading to −0.0028 volt; stimulation by λ6500 reduced the potential to 0.0016 volt or carried it back toward the initial zero point; λ5500 increased the negative potential and λ4800 carried it to −0.0038 volt, at which point the experimentation was discontinued. In curve *II* of this figure, the initial stimulus was furnished by λ5250 (green) and ultimately gave a *negative* potential increase of −0.0016 volt indicated by the point *a*. Exposure to λ6563 (red) *reversed* the potential and gave a final value of 0.00032 volt; λ4860 *reversed* it and carried it back to −0.0024 volt and so on. The points marked *d* and *e* were dual exposures to λ6000 and indicate that the second exposure to the same wave-length produced no further changes in potential. In fact the galvanometer needle remained practically quiescent during the whole of the second stimulation by λ6000.

It is of interest to note that in curve *I* λ4860 (point *2*) is practically the complement of λ5900 (point *1*) and in curve *II* λ5250 (point *a*), the complement of λ6563 (point *b*), and that in each of these cases the potential effects due to one are practically reversed and annulled by those due to the second member mentioned. The maximum potential effects were +0.0049 volt for λ5890 and −0.0016 volt for λ5250.

Curves *III* and *IV* of figure 7 are self-explanatory in view of the discussion in connection with figure 6.

We conclude, therefore, from these and other similar sets of curves obtained, that there is evidence of two processes operating in the production of retinal electrical potentials and that, in general, the shorter wave-lengths of light stimulate the production of negative potentials (from retina to nerve) while the longer wave-lengths produce positive potentials (from nerve to retina).

Complementary colors and antagonistic influences. In considering his remarks upon his researches, Gotch (4) writes:

> If it is permitted to associate these results with the well-known theories of colour vision, then it is evident that the Young-Helmholtz theory which assumes three primary colour reacting substances, red, green, violet, is in remarkable accordance with the data just put forward. In one respect only are the photo-

electric responses not in accord with that theory, for although the responses appear to show that there is a substance capable of being stimulated by darkness there is no fourth reacting substance postulated in the theory. On the other hand this is the only fact which is in any sense applicable to the rival theory advanced by Hering and even this is not in accordance with the assumption of Hering that the sensation of black, although caused by stimulation, is brought about by a change of opposite chemical type to that produced by white. The photo-electric responses, being all of the same fundamental type, are not in harmony with the assumption that there are three reacting substances, a red-green one, a blue-yellow one, and a white-black one, and that the process is of opposite sign (association or dissociation) in each member, a process of one sign occurring with red, blue or white, whilst that of the opposite sign occurs with green, yellow or black.

In his first paper Gotch (3) concludes that the results appear to indicate the localization of two photo-chemical substances in the posterior half of the eyeball, these being a substance reacting to light and a substance reacting to darkness. Each reaction, he says, is a change of the same type but for the change to occur the eye must be previously adapted, i.e., the substances must undergo phases of metabolism under conditions opposite to those which evoke the reaction effects.

Waller (7) writes:

We may see whether or no complementary colours have opposite electrical effects, and whether the excessive action of a given colour favours or disfavours the action of its complementary. . . . The effects of all sorts of colours were of the same sign. The joint effect of two complementary colours was practically the sum of their separate effects. And the prolonged excessive action of a given colour fatigued the retina to that colour just as much or as little as to the complementary colour. All colours, in fact, as regards the electrical response they elicit from the retina, give that response of the same sign and seem to act in the same direction, more or less powerfully according as they are more or less luminous.

Thus, e.g., in a given series of trials, the responses had the following relative magnitudes:

Red	4
Yellow	16
(White)	33
Green	12
Blue	9

Furthermore, all the experimental facts obtained—viz., the transitional responses from positive to negative, the immediate conversion from positive to negative by compression, the initial period of hesita-

tion—point to the conclusion that the retina, when exposed to light, is the seat of two contrary electromotive changes. And Waller says:—

> It matters little whether you imagine two single processes or one double process behind the movements of the machinery. . . . . We are all acquainted with Hering's theory of colour-vision setting forth that contrary processes of dissimilation and assimilation are aroused by complementary colours. . . . . And whether or no we happen to believe that, e.g., a red light is katabolic and a green light anabolic, we assuredly do believe in what has been termed by Bernard the axiom of general physiology—that katabolic analysis and anabolic synthesis are inconceivable apart from each other. I think that we should not hastily admit that our double electrical change is presenting us to another aspect of a familiar if somewhat nebulous colour-theory: I think we have no ground for assuming that our conclusions mean anything more than the old axiom behind a new face. A process necessitates the anti-process, and if the process is attended by an electrical change of given sign, we may expect that the anti-process will entail an electrical change of opposite sign. The chemical changes taking place in living matter are in general reversible changes.

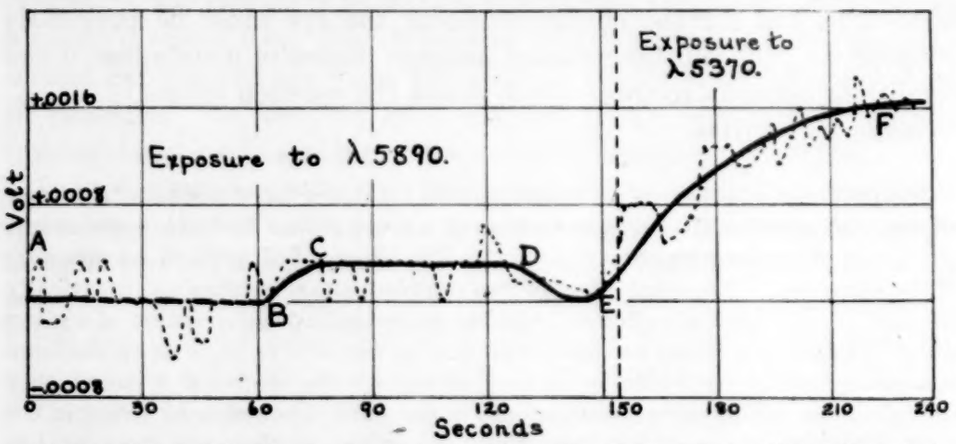

Fig. 8. Potential effects due to previous stimulation by λ4300 followed by stimulation with λ5890 and λ5370.

Sheard and McPeek (6) in two instances investigated the effects of luminous energy of a wave-length approximately complementary to the spectral region last used to stimulate the eye, under the condition that the eye had been under experimental use for some time and reached the stage when the responses to stimulation were of a minor character and did not show the pronounced and rapid changes of potential indicated in the initial stages of experimentation—such, for example, as diagrammed in figures 3 and 4. The history of the eye from which figure 8 was obtained is as follows: *a*, initial stimulation λ5900; *b*, λ4860; *c*, λ6563; *d*, λ5000; *e*, λ4300, involving a total time of 45 minutes.

Immediately previous to the obtaining of this curve the eye had been exposed to λ4300 for 4 minutes. After a period of 6 minutes in darkness it was exposed to λ5890 (sodium light) for 2½ minutes. Upon "ironing out" the irregularities, the potential-time curve has the form of the heavy line *A–B–C–D–E*. At the point *E* (after 2.5 minutes) the eye was stimulated with λ5370 (green) and the general rise of the potential positively (i.e., from nerve to retina) is indicated by the portion of the curve *E F*, with a maximum electromotive force of +0.0220 volt. The results of this experimentation indicate quite clearly that λ5376 is more nearly the complement of λ4300 than is λ5890 and that λ5370 produced a decided change in potential and in the *reverse* direction as compared with the immediately preceding stimulant, λ4300.

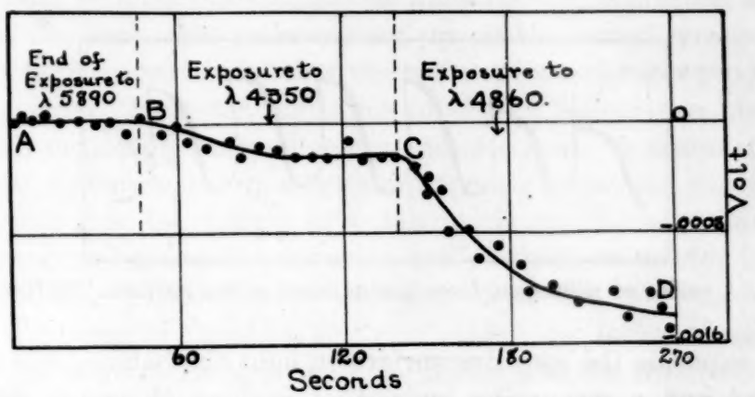

Fig. 9. Potential effects due to previous stimulation with λ5890 followed by its complementary, λ4860.

Figure 9 gives a set of conditions somewhat the opposite of those shown in figure 8. After five exposures to various portions of spectral illumination in the manner detailed in connection with the discussion of figure 8, the eyeball was exposed for 4 minutes to λ5890. The further excitation under λ4300 produced but slight effect whereas, when the pupil was illuminated with λ4860, the increase of *negative* potential (i.e., from retina to nerve) is decisively indicated in the plotted results. Hence λ4860, the complementary color to λ5890, produced a decided change in potential and in the reverse direction as compared with the immediately preceding stimulant.

These facts are further substantiated in a general way by the results shown in figure 5, in which a stimulation by λ5890 is immediately followed by its complementary color, λ4860.

ANALOGIES BETWEEN RETINAL RESPONSES AND SENSITIVE CELLS

These retinal current effects can be imitated and their counterparts observed in non-organic substances. Considerable work on this subject has been done by Bose (20). He took, for instance, a rod of silver which he beat out into the form of a hollow cup and sensitized the inside of this cup with bromine vapor. The cup was filled with water and connected through a galvanometer by non-polarizable electrodes. A current arose due to differences between the inner and outer surfaces of the cup; this was balanced by a compensating electromotive force. This gave an arrangement somewhat resembling the eye, with a sensitive layer corresponding to the retina and a less sensitive rod corresponding to the conducting nerve-stump. The apparatus, being enclosed inside a black box, was illuminated through an aperture at the

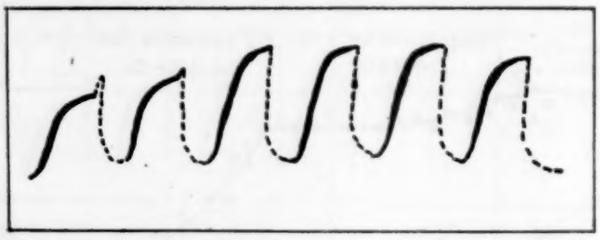

Fig. 10. A series of responses from a sensitized silver surface. (After Bose)

top; on exposing the sensitive surface to light the balance was at once destroyed and a responsive current of positive character produced. Upon cessation of light there was a fairly quick recovery. It is of interest to compare the response and recovery curves of the frog's retina as obtained by Waller and given in figure 2 with similar phenomena obtained by Bose (20) with his sensitized silver cell or other inorganic medium giving photo-electric phenomena. A sample set of data is plotted in figure 10. Waller (7) found the staircase increase which we have described elsewhere; in the record of the silver cell given in figure 10 this effect is seen at the beginning. At times in retinal responses there is an occasional slight increase of response immediately on the cessation of light after which there is the final recovery. Such an effect is occasionally encountered in the responses of the silver cell. In the responses of the retina it is found that increasing the intensity of light produces an increasing effect but the rate of increase is not uniform: the relation between stimulus and response is concave to the axis which represents the stimulus. This particular relation between stimulus

and effect is also exhibited in a remarkable manner by the sensitive cell. In résumé Bose (20, p. 167) writes:

> So we see that the response of the sensitive inorganic cell to the stimulus of light is in every way similar to that of the retina. In both we have, under normal conditions, a positive variation; in both the intensity of response up to a certain limit increases with the duration of illumination; it is affected, in both alike, by temperature; in both there is comparatively little fatigue; the increase of response with intensity of stimulus is similar in both; and finally, even in abnormalities such as reversal of response, preliminary negative twitch on commencement, and terminal, positive twitch on cessation of illumination and decline and reversal under continued action of light, parallel effects are noticed.

CONCLUDING REMARKS

The action of luminous stimulation upon the retina must lie without doubt in the ionization processes set up on the visual receptor cells. What these processes are or what the nature of the changes produced is still an enigma. However, ionization involves a dissociation into positively and negatively charged atoms and electrons. A stimulation by light must, therefore, set up a certain process of ionization whereby the positive ions may be thought of as moving in general in one direction and the negative atoms (if such exist) and electrons as moving in opposite directions. Photo-electric activities of metals, such as sodium, show a discharge of negative electrons under the action of light, the rate of discharge depending upon the metal and the character of the illumination. However, we need not think of motion of both sets of ions but of one only, for example, the electrons, and that membranes and tissues are more permeable to one than to the other of the ionized products. We may suppose that substances which in the absence of light are ionized to a degree dependent upon the equilibrium constant of the reversible dissociation action. The introduction of light into the system alters the value of this constant by increasing the rate of the dissociating change without affecting that of re-association. It seems not improbable, therefore, that the passage of one set of charged particles out of the receptor cells along the nerve trunks will thereby increase the potential difference between retina and nerve and cause a building up of a surplus of the second kind or set of charged products of the ionization process which, by accumulation, will inhibit and in large measure sweep out the oppositely charged products of dissociation when the concentration becomes sufficient.

We know that when one end of a wire is placed in contact with the cut end of a nerve and the other is touched to the uninjured nerve sheath, a current flows from the sheath to the cut portion. Hence the former is positive with respect to the latter. The Nernst-Lillie theory explains this by supposing the nerve tissues to contain an ionizable substance (or substances) the positive components of which diffuse through the sheath more readily than the negative. The sheath has, therefore, a differential permeability with respect to ions. When the cut end of the optic nerve is connected by a wire to the retina we find the retina to be more negative than the nerve end under certain conditions and *vice versa* under other conditions of stimulation. We must, therefore, it seems to us, postulate the existence of at least two ionizable substances in the receptor cells. One of these substances must give maximum ionization effects under yellow, for example, and the other similar effects under blue-violet stimulation. For one of the substances the negative ions of the receptor cells diffuse more readily than do the positive ones, while the reverse is true for the second substance.

Troland (21), independently of Nernst, makes the following definite assumptions concerning the mechanism of the visual impulse.

(1) The visual impulse consists in the actual propagation of the positive ion I_+, from the rod and cone cells along the optic nerve and tract to the cerebellum (especially to the neurons of the cuneus in the cerebral cortex). (2) This propagation takes place with the speed of the visual impulse and occurs within the neuro-fibrils which are thought of as *molecular tubes* within which it is possible for even single ions to travel without encountering great resistance. (3) The manner in which an individual ion may be imagined to be propagated is as follows: The non-fibular portion of the nerve fiber is made up of a mixture of substances, certain of which are ionized and others of which are capable of constituting an osmotic membrane which is normally equally permeable to positive and negative ions. However, when a positive ion comes into contact with one of the neuro-fibrils the surrounding neural substance acquires a slight differential permeability, so that the negative ions are capable of moving within it more readily than are the positive ions. This being the case, the loss of negative ions into the surrounding tissues—say into the myelin sheath when this is present—results in the development of a positive charge within the core. The original positive ion thus finds itself placed within the influence of a positive field. Since this is a state of disequilibrium, if the ion is free to move—and if, as will be the case, its charge is much smaller than that produced in the nerve—it will travel in one direction or the other along the neuro-fibril. If we suppose the ion to have had an original impetus in the afferent direction, it will move in this sense. The resulting process is obvious. As soon as the ion has moved into a new region of the nerve fibril the permeability of the neural substances about it for negative ions will be altered as before, a new state of disequilibrium will be produced

and the process will be repeated, the ion moving continuously in one direction within the fibril.

The general nature of the visual cerebrosis, C, follows at once from the account of the mechanism of the impulse. *The cerebral state corresponding with any condition of retinal stimulation consists simply in the presence in the cerebral cells of the specific ions which are liberated in the retina by the action of the light.*

In a recent article by Dr. H. S. Allen (22) we read:

> It is hardly necessary to assume with Helmholtz that there are three different sets of nerves having terminations in different parts of the retina. It is frequently assumed that when radiation falls upon the eye photo-chemical changes take place in certain substances in the retina, and we might imagine that there were three different chemical substances corresponding to the three primary colour sensations. But in my opinion it is unnecessary to assume that a photo-chemical change is the cause of the visual sensation. It appears to me sufficient to suppose that photo-electric action takes place in the chemical substance, or, it may be, substances, contained in the rods or cones, so that we have a separation of electrons resulting in electrification of the nerve cells which set up the nervous impulse to the sensorium. If this view is correct, the changes in the retina are identical with those which take place in a phosphorescent substance or with those which take place in the formation of a latent image on the photographic plate. After-images which are seen in the dark after viewing a brightly illuminated object are analogous to phosphorescence, and we may suppose that the excitation of the retina is due to the separation of electrons, while the after-image is due to their return. This view is supported by certain facts which are very familiar. First we know that the sensation of sight can be produced by an electrical stimulus; also it can be produced by sudden pressure. I have already referred to the fact that phosphorescence can be produced in both these ways. And in both cases we have obviously, I think, an electrical effect which may be interpreted as a separation of electrons; also it is said to be the case that the darkness-adapted eye is susceptible to Röntgen radiations. There, again, you have in all probability a process of ionization set up by the rays. I might also mention the fact that the peaked curve which shows the relation between the sensitiveness of the light of different wave-lengths bears a very close resemblance to the curve which shows the variation of photo-electric activity with wave-length.

On the other hand, Dr. W. W. Coblentz (23) of the Bureau of Standards writing on *Photo-Electric Cells and the Eye*, concludes that "In the writer's opinion, evidence is not sufficient to conclude that there is a connection between the phenomena of color perception and brightness and the phenomena of photo-electric sensitivity of inanimate material" and that "Considered as a whole, the phenomenon of spectro-photo-electric sensitivity in solids is only vaguely, if at all, similar to the visual response."

POINTS OF GENERAL AGREEMENT

From these researches one may gather the following, which are quite generally verified:

1. That positive retinal currents are manifested and positive electromotive forces established when an eye is illuminated by light.

2. That the source of these electrical changes lies in the retinal structure or in the posterior part of the eyeball.

3. That luminous stimuli of different wave-lengths—from red to violet—give initially positive retinal currents. Hence all responses are of the same general type, and act in accordance with their luminosity.

4. That the stimulus-response relation is such that the curve representing it is concave to the axis representing the stimulus.

5. That fatugue is less pronounced in the case of the retina than in that of muscle.

6. That the time-relations of the responses or periods of latency after application of stimuli are of the order of $\frac{2}{10}$ to $\frac{1}{2}$ a second.

7. That there is a characteristic response to darkness as well as to light stimulation.

8. That the range of light vibrations under which animals' (frogs quite largely) eyeballs give definite photo-electric responses corresponds very closely to the range of vision of our own color sensations.

9. That photo-electric effects with sensitive cells may be obtained which are wholly analogous to phenomena obtained with eyeballs.

10. That experimental results point to the localization of two photo-electrical substances in the posterior half of the eyeball or that the retina is the seat of a double electrical movement which may consist of duplex changes in one substance or of two changes in two different components or substances.

BIBLIOGRAPHY

(1) Edridge-Green: Ophthalmic Review, 1916.
(2) Houstoun: Proc. Roy. Soc., London, 1917, Ser. A, xcii, 11.
(3) Gotch: Journ. Physiol., 1903, xxix, 388.
(4) Gotch: Journ. Physiol., 1904, xxxi, 1.
(5) Grotthus: Lehmann Zeit. Phys. Chem., 1908, lxiv, 89.
(6) Sheard and McPeek: Amer. Journ. Physiol., 1919, xlviii.
(7) Waller: The signs of life, 1903.
(8) Einthoven and Jolly: Journ. Exper. Physiol., 1908, i.
(9) Holmgren: Untersuch. Heidelberg, ii, 81; iii, 308.
(10) Dewar and M'Kendrick: Trans. Roy. Soc., Edinburgh, 1872, xxvii.

(11) KUHNE AND STEINER: Physiol. Untersuch. Heidelberg, 1880, iii, 327; 1881, iv, 564.
(12) WALLER: Phil. Trans. Roy. Soc., B., 1901, cxciii, 123; 1901, B, cxciv, 183; Proc. Physiol. Soc., 1900.
(13) FUCHS: Pflüger's Arch., 1894, lvi, 408; 1901, lxxxiv, 425.
(14) ENGELMANN AND GRIJNS: Helmholtz. Festschrift, 1891.
(15) DEHAAS: Onderzoek. Physiol. Lab. Leiden, 2nd. ser. vi; also Dissertation, London, 1913.
(16) PIPER: Arch. f. Anat. u. Physiol., Supp. Bd., 1905, 133.
(17) ISHIHARA: Pflüger's Arch., 1906, cxiv, 569.
(18) BRÜCKE AND GARTEN: Pflüger's Arch., 1907, cxx, 290.
(19) REIDEL: Zeitschr. f. Biol., 1918, lxix, 125.
(20) BOSE: Responses in the living and non-living, 1902.
(21) TROLAND: Amer. Journ. Physiol., 1913, xxxii, 8.
(22) ALLEN: Journ. Röntgen. 1919, Soc., xv, no. 59.
(23) COBLENTZ: Amer. Journ. Physiol. Optics, 1920, i, no. 1.

EVIDENCE OF FUNCTIONAL ACTIVITY ON THE PART OF THE CAPILLARIES AND VENULES

D. R. HOOKER

From the Physiological Laboratory of Johns Hopkins University

Although studies of capillary contractility extend back for a period of fifty years almost to the time of the discovery of vasomotor nerves, and although the profound significance of the blood stream in the capillary areas for tissue nutrition was recognized at an even earlier period, capillary function is given scant if any consideration in discussion of vascular reactions today. The simple hypothesis of vascular control by means of functional activity on the part of the arterioles has been adequate to explain experimental results. Yet in recent years a considerable amount of evidence has been collected which goes to show that this conception is inadequate. First it developed that the venous bed may participate independently of the arterial function, and more recently it appears that the capillary area likewise plays a significant rôle in these reactions. This evidence has not as yet been assimilated to the point of modifying the older hypotheses. Indeed it constitutes little more than ground work: nevertheless, when gathered together, it presents a mass of data which cannot well be neglected. Basing deductions upon this newer evidence, we may well expect the future to rapidly broaden our conceptions of vascular phenomena including therein not only reactions of the muscular arteries but likewise reactions of the thin-walled veins and endothelial capillaries and venules.

When we realize that current hypotheses are the slow outgrowth of seventy years of study which started with the epochal demonstration of vasomotor nerves by Claude Bernard in 1851, it will not surprise us, even in an era of quick advance, if some time is required to root this newer work. At the present time it is constituted of isolated efforts to demonstrate independence of function by capillary and vein; as yet no effort has been made to correlate the data in a survey of the whole field.

Anatomic structure of vessels. The present paper is concerned with reviewing this work to the end that it may be available for theoretical

and practical application. To begin with, we shall consider the newer work as regards the anatomic definition of the parts with which we have to deal. For this information I am largely indebted to Dr. F. R. Sabin, who has been kind enough also to prepare the material from which the accompanying drawing was made. Older histologists recognized the essential characteristics of arteries, capillaries and veins (1). The capillary is a bare endothelial tube nicely adapted for the transfer of nutriment and waste to and from the tissue spaces which surround the functioning cells. Along these tubes course nerve fibers which present an exactly similar pattern and while contacts between these fiber-endings and the endothelial cells have not been recognized, it is a fair assumption, in view of the similarity of the nerve and vessel patterns, that such contacts do indeed occur. In other words, histological evidence accords with the belief that the vascular capillaries have a direct nerve supply. In the case of lymphatic capillaries which are structurally akin to the vascular capillaries, the evidence of a nerve supply is even more convincing (2), so that we may be satisfied that the capillary endothelium is provided with a mechanism by which it may maintain influential relations with distal parts of the body.

The endothelial layer of cells which constitutes the capillary is continued into the arteries and veins and forms the innermost layer of the intima lining these vessels. Superimposed upon this layer are tissue structures which give character to and differentiate the different vascular areas. The essential difference between the larger arteries and veins is found in the relative thickness of their walls. In the arteries smooth muscle predominates while in the veins, although a goodly supply of smooth muscle is present, white and yellow elastic tissue predominates and the total thickness of the wall is much less.

In the larger blood vessels two types of nerves are found forming plexiform nets on and in the walls. The one type, motor in character, appears to be distributed so that contact is made with each muscular element. This set of fibers undoubtedly mediates functional vasomotor responses acting upon the muscular coats. The other type, described by Dogiel as distributed to the three tunic layers of the vessel wall, terminates in a kind of flattened end-plate or specialized structure. This set of fibers thus has the character of sensory nerves. Pressure stimuli within the aorta and inferior vena cava have been shown to give rise to a reflex cardiac response. Thus Eyster and Hooker (3) found that a sudden access of pressure in an isolated loop of the thoracic aorta caused a reflex slowing of the heart. Both afferent and efferent

paths of this reflex are in part if not in whole in the vagus since the reaction occurs after section of the accelerator and depressor nerves. Bainbridge (4) has demonstrated the converse reaction, namely, that a sudden rise of caval pressure produces a reflex cardiac acceleration. The importance of the depressor nerve as a sensory path for the control of arterial tone has long been recognized. Eyster and Hooker found no evidence that these fibers arise in the heart and concluded that their distribution was confined largely to the thoracic aorta.

The point may therefore be made that the larger vessels provided with both sensory and motor innervation constitute a mechanism which, in conjunction with the heart, is able to preside over the pressure conditions of the blood as it is supplied and received from the distal and smaller branches of the vascular tree.

Between the arteries proper and the capillaries lie the arterioles, and between the capillaries and the veins proper lie the venules. Whether we should distinguish between an arterial and a venous capillary is an open question. If such a distinction is made it must rest upon functional grounds because from a structural viewpoint such a clean-cut distinction is impossible. Histologically, capillaries of different size may be recognized but they are structurally uniform in character. The terms "arterial" and "venous" in this connection imply a change in character of the contained blood. That we can distinguish in a capillary net a sharp point at which such a change occurs seems quite improbable. Variables in the tissue call for oxygen and in the rate of the blood stream would be likely to shift this point now in one direction and now in the other. The attractive hypothesis of Starling (5) that the osmotic pressure of the blood proteins serves to draw back into the vascular bed fluid which has filtered into the tissue spaces under the hydrostatic force of the blood pressure demands no structural difference in the capillary tube. The phenomenon of fluid transport depends simply upon a falling head of pressure.

While we may not therefore distinguish the capillaries as among themselves and while the older histological work offered no clean-cut distinction between arterioles and venules, newer observations have built up a sharp distinction in the latter case which is undoubtedly worthy of special emphasis because of its presumptive functional significance.

Arteriole and venule. As we pass peripheralwards along the arteries, there is a progressive decrease in the size of the lumen and in the thickness of the vessel wall. The decrease in the thickness of the vessel wall is due chiefly to the loss of connective tissue elements. The

smallest arteries are made up of the endothelial lining of the intima, a layer of connective tissue fibers and a layer of muscle fibers, all of which are continuous with the corresponding layers of the larger vessels. This transition has resulted in the elimination of the outer adventitious coat. The step down to the arteriole is accomplished by the elimination of the connective tissue layer between the endothelial lining and the muscle fibers. The arteriole, then, is composed of endothelium and muscle and is defined as that section of the vessel which has lost its intimal connective tissue and yet retains a muscular layer. As the arterioles subdivide and become smaller, the muscular layer fades away and there is left nothing but the bare endothelial tube which constitutes the capillary.

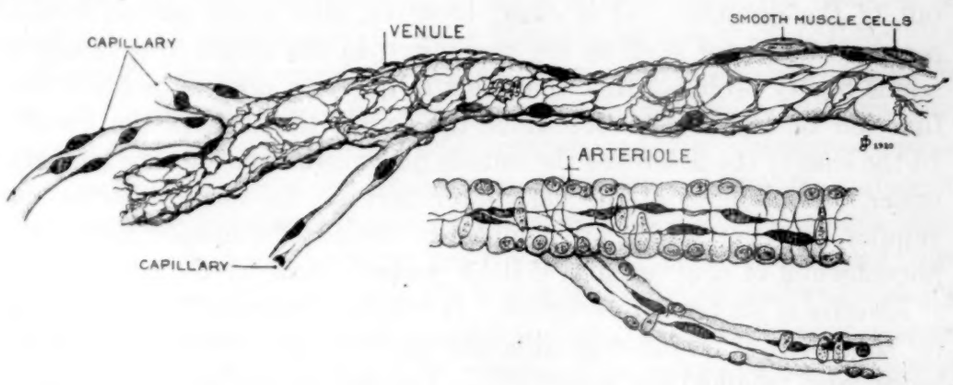

Fig. 1. Arteriole and venule from the pia mater of the pig's brain. The arteriole has a wall of endothelium and circular smooth muscle only, there being no adventitial sheath. The smaller branch shows a thinning out of the smooth muscle as the vessel approaches the capillary and the round cell near the open end of its lumen is a lymphocyte. The terminal venule has a wall of endothelium and connective tissue only, the last two longitudinal smooth muscle cells being a considerable distance from the capillaries. The capillaries are not in the pia mater but penetrate the substance of the brain.

Passing again along the capillary tubes, the latter are transformed into venules. This transition is accomplished by the addition of a connective tissue layer which takes place before there is any evidence whatsoever of the presence of muscle fibers. Thus there is a sharp structural contrast between the arteriole and the venule. The arteriole is characterized by the muscle substance intimately overlying the endothelial tube, while the venule exhibits connective tissue similarly overlying the endothelium.

This structural difference is well shown in the accompanying figure (fig. 1) drawn from preparations of the pial vessels of the brain. The

two drawings are not true to scale and the small branch given off from the vessel marked "arteriole" has not completely lost its muscular covering. In many such preparations, however, the true capillary (bare endothelial tube) is seen attached to the arteriole as it is here seen attached to the venule. The point of difference as to distribution of muscle on the arteriole and connective tissue on the venule is very clearly shown and gives an anatomic basis for the physiological results which connect the capillary and venule as one functioning structure.

Vascular nerves. There is no evidence that any but the largest vessels contain sensory nerve fibers. In the smallest vessels, arterioles, capillaries and venules, the nerve fibers are of exceedingly fine structure so that a motor and sensory differentiation on histological grounds is out of the question. It is clear, however, that there are no typical sensory end-plates such as are recognized in the larger vessels and it seems highly probable that the fibers in question are motor fibers, the function of which is to determine the contractile state of the vessels. In the case of the arterioles the muscle fibers establish the tone changes, under the influence of the vasomotor nerves. In the capillaries and venules, however, which are completely devoid of muscular tissue, the phenomenon of contractility, if it be present, must have another basis.

Theories of capillary contraction. It has been repeatedly pointed out (6) that there is no inherent difficulty in the conception that the capillaries (and venules) are contractile. Isolated protoplasmic units such as white blood cells and amoebae have the power independently to change their shape as the result of appropriate chemical stimulation. If under alterations of surface tension isolated protoplasm may change its form, the same property must adhere to the endothelial protoplasm of the capillaries. It is insufficient, however, in the case of the capillaries to conceive of a simple change of shape in their cellular constituents as a means of altering the vascular lumina without assuming that the tube length is shortened coincident with the thickening of its walls. Although the capillaries lie loosely supported by adjacent tissues and could therefore undergo change in axial length, it is more attractive to apply the other well-recognized phenomenon connected with the surface tension of cells, namely, the alteration in the plasma membrane which influences the transport of fluid, to account for capillary contractility. We may conceive then that alterations of the surface tension of the capillary endothelium, mediated by chemical or nervous influences, determines the fluid content and so the degree of turgescence of the cells.

That such a process is developed in capillary contraction is indicated by the observations of Stricker (7). Stricker, who was one of the first to interest himself in capillary function (1865), noted that when the capillaries constricted in the nictitating membrane of the frog as the result of appropriate stimulation, they disappeared almost entirely from view, to return again after the constriction passed off. This observation suggests that the distention of the cells stretching the cell membrane thins the latter out to such an extent that it becomes invisible. It is of interest to note in this connection that Stricker himself regarded the capillary contraction as associated with a turgescence of the endothelial cells.

Not all the observers of capillary phenomena have accepted the view that the capillary endothelium as such is responsible for changes in caliber of these vessels. In 1873 Rouget (8) described perivascular cells lying on and encircling the capillary wall which were regarded as the functional agents in alteration of capillary tone. It was conceived that these perivascular cells contract and in so doing accomplish a passive infolding of the capillary endothelium. Although there was no more histological evidence of the contractility of these cells than of the endothelium, the view was widely adopted. Four years before Rouget's publication, Golubew (9), who was the first to confirm Stricker's observations, had indeed suggested that certain spindle elements on the capillary walls contract into spheres, thus occluding the lumina of the vessels and so functioning as a constriction. Tarchanoff (10) adopted the same view, as did Severini (11) and Mayer (12) among the earlier observers. Among later workers Steinach and Kahn (1903) (13) alone adhere to this position.

Roy and Brown (1879) (14), among the older workers, accept Stricker's view that the endothelial cells are the prime factors in mediating changes of capillary tone. Recent workers, Dale (15), Krogh (16), Bayliss (6) and Hooker (17), adopt directly or by inference the position that the endothelial cells are alone responsible for the capillary phenomena. While this viewpoint harmonizes best with modern histological data and accords with observations of motility in the lymphatic capillaries in the tail of the living tadpole, as made by Clark (18) and others, it is to be noted that none of them worked with sufficiently high magnifications to reveal the structural details of the capillaries the function of which was under investigation.

Evidences of capillary function. Investigation of capillary function has been directed to the study of dilatation and of constriction by the

two methods of nervous and chemical stimulation. In general terms we may say that chemical stimuli may mediate both constriction and dilatation, while up to the present time nervous stimuli have been shown to cause only constriction. In the present effort to present the current viewpoint on the functional activity of these vessels, it will be expedient to emphasize the more recent contributions to the subject, making only casual reference to studies prior to 1900.

In 1903 Steinach and Kahn (13) published a paper which, although largely concerned with repetition and confirmation of earlier work, is an important contribution to the subject. Steinach and Kahn adopt the view advanced by Rouget in 1873 and maintained by S. Meyer, that certain perivascular spindle elements are responsible for "capillary constriction." The figures used by Rouget and reproduced by Steinach and Kahn to illustrate these structures suggest a striking identity with the muscular elements characteristic of the arteriole, as revealed by modern methods. This fact is apparent if one compares Rouget's figures with the drawing of Doctor Sabin's preparation given as my figure 1. One is thus led to conclude that the vessels which Steinach and Kahn regarded as capillaries and which they said contract are, according to modern histological methods, in fact arterioles and not capillaries. Steinach and Kahn do not show these cells in their drawings of contracting capillaries and it appears from the text as if the capillaries which they had under investigation showed the scattered, longitudinally-lying smooth muscle cells which are found in the area of transition from arteriole to capillary. Their drawings indicate, nevertheless, a collapse and infolding of the capillary wall in the process of contraction which would seem to be impossible of accomplishment except by some force acting from without and certainly could not be due to a turgescence of the endothelial cells.

We are consequently unable to reach a satisfactory conclusion as to the manner in which the occlusion of the capillary lumen is accomplished. As between turgescence of the endothelial cells and activity of the perivascular spindle cells of Rouget, the former is more in accord with the newer histological developments. The question is unsettled, however, and is deserving of further work.

As regards the contractility of the capillaries, however, Steinach and Kahn present positive evidence. Most of their efforts were directed to a study of the effects of electrical stimulation applied to the isolated nictitating membrane of the frog. This work confirms Stricker, Golubew, Tarchanoff and Biedl (19). They used both galvanic and faradic currents and found, in the main, galvanic stimuli more effective.

It is not easy even with the nictitating membrane of the frog to obtain both clear and irritable preparations. When so obtained, however, the application of stimuli gave sharp pictures of vascular contraction which, under favorable circumstances, might be repeated ten to twenty times. Sometimes the area of responding capillaries was extensive, sometimes only localized responses were obtained.

The amount of contraction obtained by the direct application of stimuli to the tissue is given in the accompanying table:

Measurements of capillaries at rest and during stimulation (Steinach and Kahn)

OBSERVATION	DIAMETER IN RELAXATION	DIAMETER IN CONTRACTION	CAPILLARY LUMEN	MATERIAL OBSERVED
	μ	μ		
1	26.0	12	Still patent	Nictitating membrane
2	24.0	7	Invisible	Nictitating membrane
3	23.0	15	Narrowed	Periesophageal membrane
4	22.0	6	Much narrowed	Nictitating membrane
5	21.0	4	Invisible	Nictitating membrane
6	19.0	3	Invisible	Nictitating membrane
7	18.0	8	Narrowed	Nictitating membrane
8	18.0	10	Invisible	Nictitating membrane
9	18.0	14	Notched	Nictitating membrane
10	14.0	6	Invisible	Nictitating membrane
11	10.5	3	Invisible	Nictitating membrane
12	8.4	6	Much narrowed	Nictitating membrane

The authors adduce this evidence to substantiate their contention that turgescence of the endothelium cannot explain the contraction because the *diameter of the whole capillary* is reduced. A contraction of the lumen by turgescence could not decrease the outside diameter of the vessel. But the table is also of interest in giving concrete evidence of the amount of change which may take place in capillary diameter. It shows an average decrease in diameter from 18.5 μ to 7.8 μ. Interpreting these values in terms of sectional area, we find the cross section of the relaxed capillary is 0.0011 sq. mm. as compared with 0.0002 sq. mm. for the contracted capillary, a fivefold difference. The potential efficiency of the capillary in the regulation of the distribution and pressure of the blood is thus seen to be enormous.

The time required for complete contraction is from one to three seconds. Dilatation takes an appreciably longer time. Single stimuli, even of highest intensity, are quite without effect; it requires a summating series to elicit a response.

A significant observation made in connection with the use of this isolated tissue was that the *venule* contracts as well as the capillary. The contractile process is essentially like that of the capillary but is less energetic, develops more slowly and seldom goes on to complete constriction of the lumen. The relaxation phase is described as taking a shorter time than in the capillary proper. These differences of behavior in the venules are ascribed to the fact that there are fewer contractile elements in their walls.

Results similar to the above were obtained with the membrana perioesophagealis. This tissue, however, lacks many of the advantages of the nictitating membrane and was employed relatively little. The essential results obtained with the frog were demonstrated in the use of mammalian tissue, namely, a decrease in cross section and a folding-up of the intima upon electrical stimulation. The omental vessels of the guinea pig and cat were observed. Omental tissue of young cats excised and observed in a warm moist chamber yielded the most satisfactory results. Capillaries of less than 10μ diameter were not observed to contract. But they regard the larger vessels as being true capillaries and sustaining the point. This would seem to be a wholly justified assumption. Cats' red blood cells are 6μ in diameter. A capillary tube of much less than 10μ would obviously impede the blood movement. Since they frequently noticed vessels of 10μ reduced to 3–4μ diameter under stimulation, it is a fair presumption that the smaller capillaries which failed to respond to stimulation were in fact already contracted.

Biedl, and Tarchanoff before him, had sought to demonstrate a nervous control of capillary tone without success. Steinach and Kahn, however, convinced from their results on direct stimulation of tissue that the capillaries must be under nervous control, persisted in the attack until they were able to make a preparation which gave unequivocal results. The reader must be referred to the original paper for the details. It is sufficient here to state that tracing back the nerves from the nictitating membrane, they were able to stimulate these fibers at their exit from the cord and at the same time observe the behavior of the capillaries in the nictitating membrane in situ. Under these circumstances the authors found that nerve stimulation gave as pronounced a capillary response as was obtained by direct stimulation. Here again single stimuli were ineffective and the direct current was superior in potency to induced electricity. The latency was appreciably greater than by direct stimulation, four to five seconds in favorable cases, but

sometimes twenty or more seconds. Efforts to obtain reflex capillary constriction were without avail although they were able to show a reflex effect upon small arteries close to their capillary field. In spite of their failure in this regard, they believe that reflex regulation of capillary tone will yet be demonstrated.

Do the vessels as such facilitate blood movement? One other point brought out in this paper deserves mention, namely, the fact that following nerve stimulation the capillaries are prone to exhibit rhythmic contractility. This observation is of interest in view of the question frequently raised in the literature regarding an extra cardiac force in the propulsion of blood. A recent critical review of this subject by Fleisch (20) emphasizes the improbability that the rhythmic contractions exhibited by the excised arterial strip can be interpreted as evidence of a vascular force to facilitate the blood movement. Hess (21) has shown for the pulsating veins in the bat's wing by cinematographic photographs that the time relations between contraction and relaxation are such as would facilitate blood movement. The contraction phase is about one-third and the relaxation phase about two-thirds of the cycle. Records of the arterial pulse wave show a reversed time reaction; that is, the collapsing (constriction) phase is two-thirds and the distending (relaxation) phase is one-third of the cycle. It is obvious that such a movement, even if more than passive, would not facilitate the propulsion of the blood. Furthermore, as Fleisch points out, the rate of contraction of isolated arterial preparations at body temperature is less than the heart rate. If therefore such vascular movements occurred in the normal intact animal, they would be as likely to impede as to facilitate the force of the heart beat since they are neither synchronous with it nor can they travel with the same rapidity over the vascular tube.

Fleisch thus concludes that there is no convincing evidence that the arteries furnish an extra-cardiac circulatory force. The existence of such a force in the veins is perhaps an open question, but in the case of the capillaries, since cinematographic records show no rhythmic change of caliber, the latter must also be excluded as participating circulatory aids. The most recent work makes such a sweeping statement on the capillary function very questionable. Stricker in 1865 described the observation of a varicosity on a frog's capillary which moved forward, suggesting to him peristaltic activity on the part of the endothelial tube. Also, as will be discussed more fully later, it has been shown that the mammalian capillaries and venules undergo

post-mortem changes such that they are almost completely emptied of blood shortly after death. Later they fill again and subsequently finally empty themselves. These changes take place post mortem and, since the preparation is not disturbed, cannot be passive in character both as to emptying and filling. In other words, the capillaries under these circumstances exhibit a functional response which might be of distinct moment in moving the blood under normal conditions.

It will be noted that when the blood is passed into the capillary bed its progress is slow so that to hasten it would not require a quicker action on the part of the enveloping tube than we might expect from our knowledge of the tissue concerned. Indeed, anyone who observes the capillary circulation must be struck by the varied rate of flow in this vascular field. Corpuscular stagnation in one vessel is paralleled by a rapid stream in an adjacent vessel supplied by the same arteriole. There is no doubt that some force distal to the arteriole is responsible for the condition. And this conviction is strengthened when we note that the conditions of flow in two such neighboring capillaries may be reversed without apparent change in the arteriole. This phenomenon may be especially well seen in the capillaries of the finger in the skin overlying the matrix of the finger nail (22).

This is not to say that an inherent force in the capillaries and venules actually propels the blood stream onward. Rather, we should state that the tissues involved, the capillary and venular endothelium, possess a capacity readily adaptable to such a function under the conditions which prevail. It seems therefore quite probable that an extra-cardiac force of very considerable significance to the circulation may yet be shown to be present in the capillaries and venules. Up to the present no one has attempted to show, except in the case of the bat's wing, that the veins proper facilitate blood movement by contractile activity. There is abundant evidence that the tone of the veins is an important factor in circulatory regulation (23) and we know that respiratory, joint and muscular movements facilitate by a massage effect the movement from valve segment to valve segment of the blood in the large veins. The sole indication, however, that the veins may move the blood by a kind of peristaltic activity is found in analogous observations that the large lymphatic trunks in situ may exhibit spontaneous rhythmicity of a peristaltic character. Gunn and Chavasse (24) found that the application of epinephrin to the inferior vena cava at a considerable distance from the heart elicited rhythmic contractions in a formerly quiet preparation, but it may be doubted if the extension of the auricular

function out over the great veins is to be regarded as a functional aid to the movement of the blood.

Chemical stimuli. I have dwelt at considerable length on the work of Steinach and Kahn because it presents in large part the viewpoint, methods and results of earlier investigators. It must be pointed out, however, that effort was not limited to a study of electrical stimulation. Thus Stricker used chemical and mechanical stimulation as well and obtained excellent results on the capillaries of the tail of the living tadpole. The results with ammonia vapor were particularly striking. When the tissue was exposed to the vapor for three or four seconds and then examined under the microscope, Stricker saw the capillary lumina almost disappear and then dilate wide enough for a corpuscle to pass. This phenomenon occurred about twice in fifteen minutes. Similarly Tarchanoff made use of alcohol, ether, ammonia, ferric chloride, acetic acid and also heat and obtained capillary constriction. The character of the functional response of the vessels is not made very clear and it appears that the vessels very often failed to dilate after the stimulus was removed. Roy and Brown in 1879 and also Lister (25) in 1858 found that irritant chemical substances produced capillary dilatation. Lister considered this response secondary to arteriolar dilatation, but Roy and Brown showed conclusively by concomitant pressure determinations that the effect was directly upon the capillaries. Biedl studied the effect of heat. He reports seeing all the peripheral vessels of the frog's mesentery (arterioles, capillaries and venules) contract on the application of salt solution at 45°C. and dilate again when the heat was removed.

Severini (11), as early as 1878, published a monographic study of capillary function. In this he states that the application of oxygen causes the capillaries to contract, while the application of CO_2 causes them to relax. These observations were made upon the frog and also upon the mesentery of the guinea pig. Lister, likewise, had observed the effect of CO_2 and used the phenomenon to explain the congestion of the lungs in asphyxia. These results could not be confirmed either by Tarchanoff or by Roy (26), but are nevertheless of interest in being the first definite suggestion of the importance of these gases in the regulation of vascular tone. Quite recently, however, Krogh has been able to confirm them for carbonic acid.

Mechanical stimuli. When a blunt-pointed object is drawn across the skin with some pressure, a vascular skin reaction is developed which, in certain circumstances, is regarded clinically as symptomatic of

vasomotor instability. The reaction was first described by Marey in 1858 (27), who regarded it as dependent upon a capillary response. The track of the object is first white, due to expression of the underlying capillary blood. The white line turns red in a few seconds and may become bordered with white bands, with flushed bands or with both white and flushed bands, the white bands lying next the red line. The reaction persists for varying lengths of time, usually a matter of minutes. The duration and brilliance of the phenomenon depend upon the sensitivity of the vessels and in certain cases may go on to the development of definite urticarial wheals (urticaria factitia).

This reaction is generally known as dermatographism and in extreme form is regarded as an angioneurosis. Müller (28), in a discussion of the phenomenon and its diagnostic significance, published some beautiful photographs indicative of the character of responses obtained in different clinical conditions. Ryan (29) has sought to apply the test as a criterion of physiological fatigue. The latter's results indicate that the primary white line fades more quickly when the subject is fatigued or suffers from an infectious disease.

In 1917 Cotton, Slade and Lewis (30) attacked the subject of dermatographia as exhibited in soldiers with "irritable heart," and present evidence to show that some of the conditions observed can best be explained on the assumption of contractility in the capillaries. They regard the red line of the tâche and the flush which may surround it as due to arteriolar dilatation which floods more blood into the capillary bed. The white tâche and the white area beside a red tâche, however, are accepted as due to capillary constriction. When a red tâche has been established its tint is first that of the skin, but blueness is rapidly developed, they argue, by capillary dilatation which facilitates oxygen diffusion and in the absence of increased blood flow develops the blue color.

The most convincing evidence adduced by these investigators for capillary activity is found in experiments in which the blood flow in an arm was completely shut off by a supra-systolic pressure in a sphygmomanometer cuff. When the arm was held horizontal at the time this pressure was applied, the color of the skin remained normal and the stroke of a blunt instrument developed a white tâche just as distinct as under normal circulatory conditions. Since this reaction could be obtained at any time during ten minutes of applied pressure, they quite properly assume a stabilized pressure in the vessels. If under such circumstances the capillaries empty themselves of blood on appropriate

stimulation, the result cannot be ascribed to arteriolar constriction and must be due to a response on the part of the capillaries themselves. Under the same conditions stronger mechanical stimulation of the skin produced a red tâche which was assumed to be due to a capillary dilatation, thereby permitting blood to enter from the larger neighboring vessels.

Following the same procedure as regards occluding the larger vessels, these authors find that the cutaneous injection of a few drops of 1:30,000 adrenalin causes a localized pallor much as is the case when the substance is injected with circulation intact. Pallor in the latter case is accepted as due to constriction of the arterioles and this the authors do not deny. But when the blood flow in the large vessels is shut off, the pallor must be due to an effect upon the capillaries directly. We shall have more to say of the effect of epinephrin in later pages.

Sollmann (44) has studied the reactions produced by the application of various substances to the scarified skin. Urticarial reactions were obtained with histamine, peptone, morphine, formaldehyde, atropin, cocaine, calcium chloride and urea, while blanching occurred with epinephrin, pituitary and hydrastin. Although localized flushing or blanching of the skin was conspicuous in these reactions, Sollmann was unable to connect them with any known chemic or pharmacologic relations because other substances having equally potent vasomotor properties were without effect. He ascribed the urticarial responses to a specific increase in capillary permeability. It would be of interest with our present methods and knowledge of capillary function to again study these reactions to observe directly the behavior of the capillaries and venules.

Briscoe (31) made a careful study of the venous and capillary pressures in a number of individuals suffering from the condition known as "irritable heart" who exhibited local asphyxia or cyanosis of the hands. She made use of the apparatus described by Hooker (32) and used as criteria for the venous pressure the collapse of the superficial vein, and for the capillary pressure the blanching of the skin under the glass capsule. Throughout her observations the venous pressure was found within normal limits (average 11.4 cm. H_2O). The capillary pressure was, however, uniformly high and sometimes as high as twice the normal. The average capillary pressure for the controls was 23.5 cm. H_2O, and for the subjects with well-marked blue hands 36.9 cm. H_2O. In these cases there is little doubt of capillary stagnation; the blue color, the tendency to perspire, and the dark color of the blood

which exudes from a skin prick, all point to this condition. The stasis might be due to arteriolar constriction. This factor is, however, excluded by the high capillary pressure. Hence the only mechanism which could develop a high capillary pressure with engorgement accompanied by a normal venous pressure would be a constriction of the venules and this is the explanation of the condition which Briscoe accepts.

We find in this explanation the first suggestion of independent function on the part of the venules. I shall present direct evidence of the ability of these vessels to contract and to dilate, and it is quite conceivable that they may do so in these cases of local cyanosis although Briscoe's evidence is the sole indication of it at the present time. It must be remembered also that the criterion of color change in the skin for capillary pressure is open to serious objections, as has been pointed out by Danzer and Hooker, and especially on a cyanotic skin this criterion must be difficult of interpretation.

Returning again to the results of direct experimentation of capillary function, we have the recent contributions of Krogh (1919). Krogh reached the conclusion that in muscle at rest many capillary channels are occluded, and that as the result of activity the number of patent capillaries is enormously increased. This conclusion rested on the study of both living and fixed specimens.

It may be assumed that the response is due primarily to the stimulus exerted upon the capillary endothelium by the katabolic products of muscular activity. In this paper Krogh well emphasizes the inadequacy of the older concept of capillary function that the number of functioning capillaries in any tissue depended upon the tension of blood in the arterioles, and that as this tension rises more and more capillaries open up as if their tonicity varied in degree from capillary to capillary. He points out that if such an arrangement existed, the blood would always tend, when the pressure was low, to flow in certain fixed channels so that some cells would be richly supplied with oxygen while others would be in constant danger of suffering oxygen want. There is today much reason to believe that the capillaries are readily susceptible to chemical stimulation especially, and that thereby the blood, without recourse to changes in arteriolar pressure, may be shunted from one set of channels to another according to the tissue needs of their environing cells. Roy and Brown indeed noted (1879) that the caliber of capillaries is constantly changing, and that these changes are independent of the pressure of the blood supplied to them. As a matter of fact, one may

note in the skin overlying the matrix of the finger nail, when the tissue is made transparent to microscopic vision by a drop of oil, what Krogh observed in the tongue of the frog, namely, the not infrequent alternation of stasis and flow in the capillaries under observation. Danzer and Hooker found that one of the annoyances in determinations of capillary blood pressure in man, when the criterion employed was the cessation of corpuscular flow, lay in the fact that often the large and conspicuous capillaries were filled with stagnated corpuscles, and they noted that in the course of observations it was not uncommon to find that a previously stagnated vessel had become patent and that on the other hand patent vessels became stagnated. Furthermore, they found under varying experimental conditions that the pressure might rise or fall in one capillary while the pressure in its neighbor remained constant or changed in an opposite direction. The structural conditions of the vascular bed are undoubtedly such that in the vast majority of instances neighboring capillaries are supplied by the same arteriole. Variations in neighboring capillaries such as have been noted therefore must be due to inherent and independent changes in the vessels themselves.

In another place Krogh has presented further evidence of independent contractility in the capillaries of the frog as induced by both chemical and mechanical stimulation. As regards the latter, he states that punctate stimulation with a needle point may be seen to cause dilatation of a single capillary or even a short section of a single capillary. It is interesting that when such a reaction is induced, the dilatation starts at the venous end of the tube so that as the vessel fills the blood appears to flow toward the heart. When the dilatation reaches the arteriole, there is a sudden reversal to normal in the direction of the flow. I have myself repeatedly seen this reversal of flow under somewhat different conditions in the capillaries of the cat's ear. The observation that cocaine abolished the response to mechanical as well as to certain chemical stimuli suggests a nervous control of some kind. Acids, however, among the chemical stimuli, were exceptions in that cocaine did not abolish their effect. Immediately after nerve section the response to both mechanical and chemical stimulation is unaltered. With the establishment of degeneration the area of response becomes constricted until finally it is sharply localized to the point directly stimulated. These findings led Krough to the conviction that capillary tone is not only independent of the blood pressure but that the spreading of the effects of stimuli is due to axone reflexes which are

probably conveyed along sensory fibers. This position accords with the conception of Bruce (33) and of Bardy (34) that inflammatory processes resulting from the application of local irritants are essentially due to axone reflexes.

All the chemical substances investigated by Krogh (acetic and carbonic acid, adrenalin, iodine, urethane, cocaine, nicotin and amyl nitrite), when applied directly to the vascular area under observation (the frog's tongue) produced capillary dilatation. The same result followed from the application of heat, of cold and of mechanical stimulation. Electrical stimulation of the lingual nerve was without effect. These results are not in accord with the findings of other observers, as will be noted in the table which follows (table 1), nor can they be readily harmonized among themselves, i.e., the effects of heat and cold are the same. Krogh is of the opinion that capillary tonus is not of nervous origin but must depend upon some constituent of the blood, and this constituent is not oxygen.

In 1918 Hooker (35) showed that it is possible to demonstrate an appreciable rise of venous pressure on appropriate stimulation in a preparation when the arterial pressure was zero. The preparation consisted of the sigmoidal area of the large intestine in dogs, isolated except for its nervous connections via the inferior mesenteric ganglion. After the blood was washed out by perfusion, the artery was left open and the drainage vein connected with a water manometer. Direct electrical stimulation of the nerve to the part gave a rise of pressure of 7.5 cm. of water, indirect stimulation (asphyxiation) acting by way of the medulla gave a rise of 4.5 cm. of water. The possibility that the response obtained was due to activity of the intestinal musculature was excluded by coincident records of the pressure within the lumen of the gut itself. The results thus obtained were regarded as substantiating the existence of a central as well as a peripheral venopressor mechanism. Whether the contraction was in the larger vein, in the venule or in the capillary, or in all of them together, was not possible of determination, but it was believed that the arteriole and artery were effectively excluded. In our present knowledge we should probably emphasize the participation of the capillary and venule although the direct evidence is not at hand in this particular experiment.

In the hands of the earlier observers chemical stimuli in the main produced capillary constriction. Later work emphasizes dilatation as the result of such procedure. The following table summarizes the substances employed, the tissues used, the results obtained and the names of the observers.

TABLE 1

To show the results obtained on capillaries by various observers by the use of different chemical substances

SUBSTANCE EMPLOYED	TISSUES OBSERVED	RESULTS OBTAINED	OBSERVERS
Chloroform	Frog's web (intact)	Dilatation	Lister
	Frog's mesentery (intact)	Dilatation	Roy and Brown
Ammonia	Frog's nictitating membrane (isolated)	Constriction	Stricker
	Tadpole's tail (intact and isolated)	Constriction	Stricker
	Tadpole's tail (intact)	Constriction	Tarchanoff
Alkali	Frog's nictitating membrane (isolated)	Constriction	Stricker
	Tadpole's tail (intact and isolated)	Constriction	Stricker
Acetic acid	Frog's nictitating membrane (isolated)	Constriction	Golubew
	Frog's nictitating membrane (isolated)	Constriction	Tarchanoff
	Frog's mucous membrane of mouth (isolated)	Constriction	Golubew
	Frog's muscle (isolated)	Constriction	Golubew
	Frog's skin (isolated)	Constriction	Golubew
	Tadpole's tail (isolated)	Constriction	Golubew
	Tadpole's tail (intact)	Constriction	Tarchanoff
	Frog's web (isolated)	Constriction	Golubew
	Frog's web (isolated and intact)	Constriction	Tarchanoff
	Frog's tongue (intact)	Dilatation	Krogh
	Frog's web (intact)	Dilatation	Lister
Carbon dioxide	Guinea pig's mesentery	Dilatation	Severini
	Frog's tongue (intact)	Dilatation	Krogh
	Frog's web (intact)	Dilatation	
Alcohol, ether, ferric chloride	Tadpole's tail (intact)	Constriction	Tarchanoff
	Frog's nictitating membrane (isolated)	Constriction	Tarchanoff
	Frog's web (isolated and intact)	Constriction	Tarchanoff
Heat	Frog's web (isolated and intact)	Constriction	Tarchanoff
	Frog's mesentery (intact)	Constriction	Biedl
	Frog's tongue (intact)	Dilatation	Krogh
	Frog's web (intact)	Dilatation	Lister
Cold	Frog's tongue (intact)	Dilatation	Krogh

TABLE 1—*Concluded*

SUBSTANCE EMPLOYED	TISSUES OBSERVED	RESULTS OBTAINED	OBSERVERS
Oxygen	Guinea pig's mesentery	Constriction	Severini
Gold salts, etc.	Frog's mesentery (intact)	Dilatation	Heubner
Adrenalin	Frog's tongue (intact)	Dilatation	Krogh
	Human skin (intact)	Constriction	Cotton, Slade and Lewis
	Cat's ear (intact)	Constriction	Hooker
Histamine	Human skin (intact)	Dilatation	Sollmann
	Cat's mesentery (intact)	Dilatation	Rich
	Cat's ear (intact)	Dilatation	Hooker
Mustard	Frog's web (intact)	Dilatation	Lister
Oil of turpentine	Frog's web (intact)	Dilatation	Lister
Cantharides	Frog's web (intact)	Dilatation	Lister
Croton oil	Frog's web (intact)	Dilatation	Lister
Iodine	Frog's web (intact)	Dilatation	Lister
	Frog's tongue (intact)	Dilatation	Krogh
Urethane	Frog's tongue (intact)	Dilatation	Krogh
Cocaine	Frog's tongue (intact)	Dilatation	Krogh
Nicotine	Frog's tongue (intact)	Dilatation	Krogh
Amyl nitrite	Frog's tongue (intact)	Dilatation	Krogh

The above table comprises the effects of various substances, most of which have an inconspicuous, evanescent effect. In the case of two of them, however, gold salts and histamine, the effect is striking and permanent. Heubner (36) especially has shown that the intravenous injection of the double chloride of gold and sodium produces a shock-like prostration accompanied by a remarkable engorgement and stasis of the capillaries and veins. Post-mortem examination showed, in the case of mammals, distended and conspicuous veins which could be traced as fine twigs onto the surface of the intestine. The peritoneal surfaces were injected and the mucous membrane had a similar hyperemic appearance. Microscopically all the tissue exhibited an extensive dilatation of the venules and capillaries, and not infrequently ecchymotic areas. The small arteries were everywhere strongly contracted, usually with completely occluded lumina.

In frogs Heubner was able to observe a picture much like the above develop under the microscope. Curarized frogs were prepared to observe the mesenteric circulation and doses of the double gold salt equivalent to 0.25 mgm. metallic gold were injected intravenously. During the injection the vascular picture remained unchanged. In a minute or less after the injection there developed a sudden and extreme increase in the number of capillaries visible and at the same time the intestine assumed an intensely red color. Shortly after this change the circulation came to a complete standstill. These results indicate that the substance used is an endothelial poison which primarily attacks the capillaries and venules. The dilatation and engorgement of the large veins suggest also that these vessels are affected. How far the effects observed are dependent on the central nervous system is an open question. And it seems quite improbable, as Heubner points out, that the picture is developed simply as a result of arteriolar dilatation since vasomotor paralysis, so far as is known, does not establish a hyperemia with a tendency to ecchymosis.

A similar type of reaction has been predicated by Dale (15) and his co-workers for histamine. It is true that these investigators were unable to present direct evidence of the toxic action of histamine on capillary endothelium; nevertheless their indirect evidence is so strong that one does not hesitate to accept their conclusions.

Histamine (β-Iminazolylethylamine) is a substance first isolated from ergot but subsequently prepared by decarboxylation of histadine and now believed by some to occur in various animal tissues (37). It exhibits the paradoxical action of intensively stimulating smooth muscle and yet, when introduced into the blood stream of cats and dogs, of causing a profound and fatal depression of the arterial blood pressure. The paradox is accentuated by the fact that isolated arterial strips, like smooth muscle derived from the uterus and other parts of the body, when brought into contact with solutions containing this substance, respond with powerful contraction. The signs and symptoms presented by an anesthetized animal acutely poisoned with histamine (about 6 mgm. per kilo intravenously) resemble in many respects the condition of traumatic shock. Dale and his co-workers, by reasoning and experiments, reached the conclusion without observation of the vessels concerned, that histamine is a capillary poison producing an effect upon the general circulation analogous to that established by the injection of gold salts. Cannon (38) had previously presented evidence that the capillaries are packed with red cells in wound shock and postu-

lated this condition of "exemia" as the causative factor in the condition. This conclusion has since been confirmed by Rich (39) and by the present writer (17). Rich developed a method by which he could obtain instant fixation of the omental tissues in situ. He was then able to study the vascular picture under the microscope after the omentum had been removed and spread upon a slide. Preparations made in this manner at various times after the intravenous injection of histamine in cats and controlled by specimens in which physiological salt solution was injected gave clear evidence that an immediate effect of the histamine was to increase greatly the number and size of the capillaries.

Similarly I have been able to show by photographs of the capillaries in the cat's ear that here too histamine causes a profound dilatation of the capillaries and venules. The vessels of this area may be less sensitive than in the omentum; at any rate I frequently observed that the dilatation due to histamine was preceded by a period of constriction. This constriction was usually fleeting in character and while of considerable moment in regard to capillary function is not to be regarded as disputing the hypothesis of histamine action under discussion. In connection with my observation that histamine dilates the venules as well as the capillaries, it is of interest to point out that measurement of the venous pressure in these experiments showed a fall of venous pressure corresponding in magnitude to the drop in arterial pressure. Dale states that the portal venous system is collapsed and empty. It follows therefore that the dilatation of capillaries and venules must be able to accommodate a very considerable amount of blood since the pressure both in artery and vein falls too quickly to be wholly accounted for by transudation of plasma into the tissue spaces.

Histamine, as Dale pointed out, does not abolish arterial vasoconstriction upon sensory nerve stimulation. It does, however, completely depress the response of capillary and venule to peripheral nerve stimulation, as will shortly be emphasized, indicating again its potency as a poison for the capillary endothelium.

The present writer has found that, using Lombard's method (40) of a drop of oil upon the skin, the cat and other animals can be used advantageously in studying the capillary circulation. After the inner surface of the ear of a cat with non-pigmented skin has been shaved, cleaned and dried, the pinna is spread out and sealed to the board so that its surface is flat and slightly below the level of the heart. A drop of castor oil is flooded over the area to be studied and with direct illumination the microscope reveals an exquisite capillary net about

the hair follicles, in which corpuscular flow can be readily visualized. In addition to the capillaries, venules can be seen but, in the ear, the arterioles are not visible presumably because they lie deeper in the tissue. The nictitating membrane gives an even more beautiful vascular picture, showing arterioles, capillaries and venules, but is not adapted for experimental procedure because in the anesthetized living animal it cannot readily be brought below the heart level and studied. The level of the vessels with respect to the heart is of prime importance because, using low power magnification (70 ×), the vessels are not visualized as such but by their contained blood. Thus the capillaries and venules are assumed to contract when they empty of blood. When these vessels are below the heart level they are subject to a hydrostatic column of venous blood. Constriction of the arterioles alone, therefore, could not empty them, and if they do empty or appreciably change their size as the result of experimental procedure, the change must be due to inherent activity in the vessels themselves.

Objection may be made to this assumption on the ground that arterial constriction could block the entrance of red blood cells into the capillaries and still allow the passage of plasma, the effect of which would be to wash out any red cells lying in the capillary channels. Lister, who was one of the first to study capillary function, indeed made this objection to one of his own experiments on the frog in 1858. In the present case, however, the objection is not valid, as will be made clear by what follows.

When an anesthetized cat is placed on a flat animal holder, the capillaries of the ear lie about 2 cm. below the heart level and are therefore subject to a venous hydrostatic pressure of 2 cm. of blood. Under these conditions, if the arterial supply to the vascular area under observation is mechanically shut off so that there is no corpuscular movement in the capillaries, the latter vessels do not empty. If now, instead of occluding the artery, the animal be promptly killed by an overdose of anesthetic, in the course of a few minutes the capillaries and venules will be emptied almost completely of blood. This emptying takes place by a movement of blood from the arterial to the venous side of the capillary net. After some time, fifteen minutes or more, the vessels again fill and it may be seen that they fill from the venous side of the bed. After a further period of time roughly coincident with the onset of skeletal muscle rigor, the capillaries and venules again empty and thereafter remain empty.

This observation, subject to certain exceptions which need not be considered here, may be readily repeated. It shows in the first place that the capillaries and venules function independently of arterial pressure changes, as was first clearly emphasized by Roy and Brown; and in the second place it shows that the removal of corpuscles from these vessels is in no sense necessarily dependent upon their being washed out by a stream of plasma.

In the research now under discussion certain other facts were presented which are of importance to our main thesis and which for the first time correlate and extend the earlier observations on the frog to the mammal. The post-mortem movement of the blood in the capillaries and venules under the conditions as defined, is clear evidence of the functional activity of these vessels. In the tip of the ear, composed of skin with a few scattered hair follicles overlying a cartilaginous base, and placed below the heart level, the only force which can cause such corpuscular movement must be in the endothelium of the capillaries and venules. The development of a similar corpuscular movement by other procedures must likewise be ascribed to endothelial activity. Of the two post-mortem capillary constrictions above described, the first is in all probability asphyxial in origin and may contribute in no small degree to the well-recognized asphyxial rise of arterial and venous blood pressure. So ready an explanation of the second constriction is not at hand. Its onset coincident with the development of skeletal muscle rigor suggests the possibility of a relationship to some similar change in endothelium. There is, however, at the present time, no scientific basis for such a suggestion.

Of foremost importance to the theory of capillary function is the effect produced by nerve stimulation. If the cervical sympathetic be sectioned and stimulated in the cat, there occurs first a cessation of corpuscular flow due, it may be assumed, to arteriolar constriction. Then the corpuscles begin to clump and to move forward in an irregular manner so that in the relatively brief time that the stimulus (electrical) is continued, the field of observation is wiped quite clean of corpuscles. Promptly after the cessation of the stimulus corpuscles appear again from the arterial side and the flow is even more abundant than before the stimulation. This result can be repeated many times on the same preparation without any indication of fatigue, and appears to be conclusive evidence that the capillaries and venules are subject to nervous control. Figure 2 reproduced the photographs taken in such an experiment.

While it was not shown in this investigation that section of the cervical sympathetic in the anesthetized cat caused new capillaries to open up or the capillaries already present to appreciably dilate, it was observed that subsequent to the nerve stimulation there was a very definite increase in capillary size. This result is suggested in the above figure but owing to the uncertainty of focus the size of vessels so reproduced is not wholly trustworthy. Photographs are to be relied upon

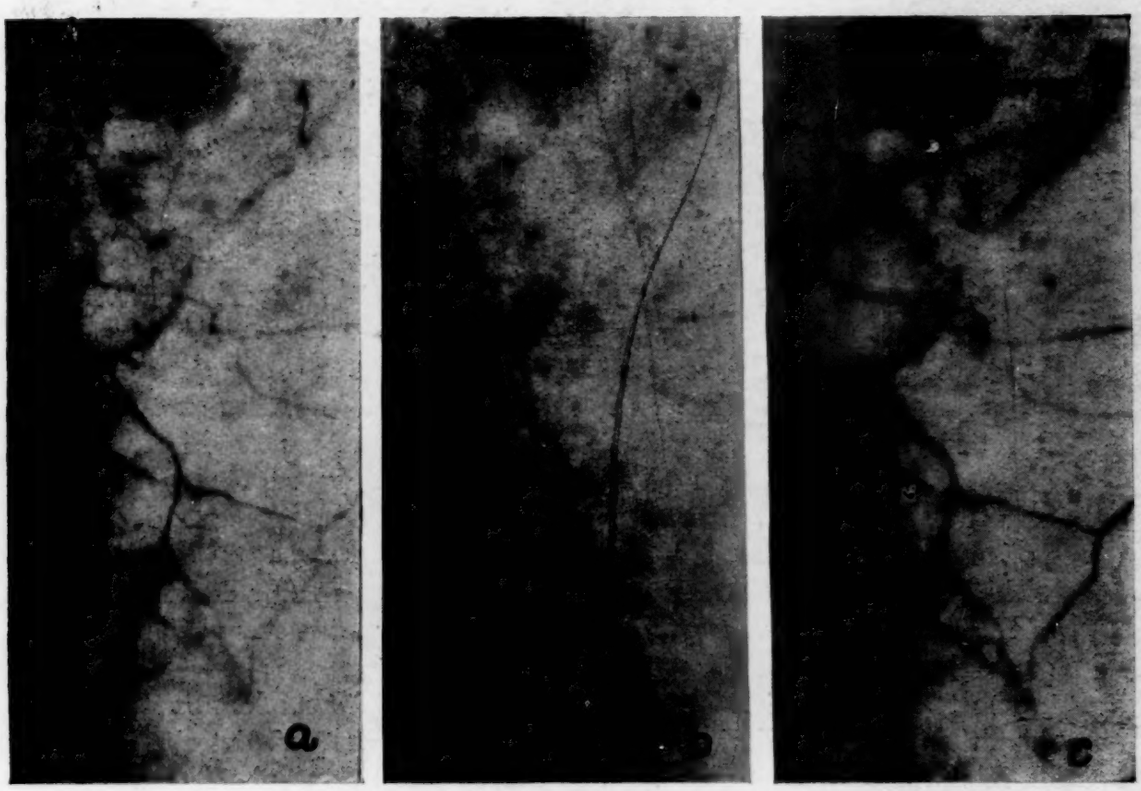

Fig. 2. Constriction of capillaries and venules by electrical stimulation of the cervical sympathetic. Magnification 90 ×. Cat *a*, before; *b*, during; and *c*, after stimulation. The negatives have not been retouched.

only when the vessels or sections of them are present or absent or when the change in the size is very conspicuous. Nevertheless it is readily possible to see by direct observation the changes above described. Further support for the contention that nerve impulses play upon and influence capillary tone is found in the action of epinephrin. Epinephrin is today generally accepted as acting only on tissues with sympathetic innervation. If 3 cc. of 1:50,000 adrenalin solution be

injected intravenously, the capillary bed behaves much as it does in the case of nerve stimulation, that is to say, the corpuscles stop, clump and then slowly and irregularly move out of the field. There can be little doubt, especially if the process is observed directly with the eye, that the events here described are directly dependent upon a functional response of the endothelium.

The above result with epinephrin is not in agreement with Krogh's experiments on the frog. Krogh found that the application of 0.1 per cent adrenalin to the surface of the tongue exposed for view under the microscope produced hyperemia. This capillary dilatation sometimes occurred independently of arterial response but usually the smaller arteries dilated also. This dilatation of the smaller arteries suggests that the effective concentration might have corresponded to the usual dilator dose. Since Krogh was able to establish this capillary effect as independent of arterial change, it may be surmised that the substance produces a reaction in the capillaries and venules in proportion to its strength, as is the case of its effect on the systemic blood pressure.

It is of considerable interest that these results also confirm Dale's hypothesis that histamine is a capillary poison. The intravenous injection of histamine (ergamine phosphate, 6 mgm. per kilo) in the anesthetized cat causes a sudden fall in arterial and venous blood pressure from which recovery may or may not take place. Direct observation of the capillaries, as well as photographs, shows that coincident with these blood presssure changes there is a corresponding dilatation of the capillaries and venules. Rich has shown the same change in the omental vessels by another method.

The toxic effect of this substance is emphasized not alone in the dilatation of the capillaries and venules which it produces, but also in its inhibition of the normal vascular responses. In an animal which has been killed by a lethal dose of histamine, the capillaries and venules are promptly dilated and remain so indefinitely. Thus they fail to show the primary constriction which has been described above as asphyxial in origin, or the secondary constriction which is, under ether death, coincident with the onset of skeletal muscle rigor, although animals killed with histamine develop skeletal rigor as readily as do the others. Similarly the injection of histamine develops such a condition in the capillaries and venules that nerve stimulation and the intravenous injection of epinephrin are entirely without effect although it may be easily shown, as was pointed out by Dale, that epinephrin injection and the stimulation of sensory nerves elicit an arterial blood

pressure rise comparable to that obtained before the injection of histamine. This arterial blood pressure rise after histamine, due to the constriction of the arterioles, since it does not affect the capillaries and venules, justifies again the assertion that when the capillaries and venules do respond to experimental procedure such response must be an active one and not passive to changes in the caliber of the arterioles and therefore in the pressure and volume of blood supplied to them.

In the foregoing pages we have sketched the development of our knowledge of the function of the capillaries and venules, and find that today we have a considerable mass of evidence indicating that the behavior of these vessels must be of very great functional significance although their coöperation in vascular reflexes has not hitherto been considered. The work of Roy and Brown, and quite recently that of Krogh and of Danzer and Hooker, has indicated the ability of the capillary bed to respond to local tissue needs by dilatation when tissue conditions tend toward asphyxiation, by constriction when such local needs have been satisfied. Thus there is a constant play of function opening up and closing various capillary channels in response to the metabolic state of the environment. When the capillary lumen is closed it may be quite empty of corpuscles, as noted by Krogh for resting muscle, or it may be packed with stagnated corpuscles as noted by Danzer and Hooker for the cutaneous capillaries of man. One would expect therefore that such changes, while ordinarily balancing one another in different parts of the body, might under special conditions exert a profound effect both upon the distribution of the whole blood and upon its partition between corpuscular elements and plasma. Cannon and Dale have emphasized the fundamental significance of the capillary bed in different forms of "shock," and Thaysen (41), in a communication which has not been available to the writer, has ascribed a remarkable variation in the red cell count in a case of polycythemia to alterations in the tone of the capillaries and precapillaries of the skin. Thaysen is reported to have observed the red cell count to rise from 5.3 million to 10.6 million in a period of twelve hours, a reaction not unlike that found by Lamson (42) to follow the injection of massive doses of epinephrin in experimental animals. Lamson has located the polycythemic reaction to epinephrin in the liver and inclines to the view that the increased red cell counts found after epinephrin, exercise and other experimental procedures are due to blood concentration rather than to the washing into the circulation of cells stored in quiescent vascular areas, as suggested by Snyder and Havens (43).

In either case the capillary bed appears responsible for the changes observed.

We have seen that the capillary bed is responsive to both chemical and nervous influences. Broadly speaking, we may say that chemical factors, so far as they have been studied, mediate dilatation of the capillaries and venules, while nerve stimulation mediates constriction of these vessels. Both these factors have long been recognized as participating in vascular reactions, but only in recent times have they come to be specifically located in their effects upon the capillary and venule, that is, upon that section of the vascular bed which is composed of bare endothelium. We may believe that the chemical regulation is usually local in character, producing changes in accordance with the passing needs of particular tissues, and that only in special or pathological conditions does the reaction extend to the body as a whole. In accordance with this view we might regard the nervous control as a force tending to restrict the capillary beds over the body as a whole, thus maintaining a tone to be played upon by chemical factors. But the story is far from complete and much remains to be done. This much, however, appears true, that we must reorganize our conception of the peripheral resistance. Hitherto the functional peripheral resistance has been limited to the arteriole, the caliber of which is controlled by a muscular coat subject to nervous and chemical influences. The resistance to the flow of blood by the capillary and venule has been regarded as essentially fixed. Now we know that beyond the arteriole the capillary and venule function actively and must therefore participate directly in vascular reactions. Whether it will be shown that these several areas of the peripheral vascular bed function independently of one another is an interesting question for the future.

BIBLIOGRAPHY

(1) Mann: Blood vessels. Quain's Anatomy, London, 1912, ii, part i, 328.
(2) Sabin: The origin and development of the lymphatic system. Johns Hopkins Hosp. Repts., 1916, xvii, 347.
(3) Eyster and Hooker: Direct and reflex response of the cardio-inhibitory center to increased blood pressure. Amer. Journ. Physiol., 1908, xxi, 373.
(4) Bainbridge: The influence of venous filling upon the rate of the heart. Journ. Physiol., 1915, l, 65.
(5) Starling: Human physiology, Philadelphia, 1920, 1069.
(6) Bayliss: Capillary circulation. Science Progress, 1919–20, xiv, 272.

(7) STRICKER: Vorlesungen über die allgemeine und experimentelle Pathologie, Wien, 1883, 675.
(8) ROUGET: Memoire sur le Developpment, la Structure et les Proprietes Physiologiques des Capillaries Sanguins et Lymphatiques. Arch. d. Physiol. Normal et Path., 1873, v, 603.
(9) GOLUBEW: Beiträge zur Kenntnis des Baues und der Entwicklungsgeschichte der Capillargefässe des Frosches. Arch. f. mikros. Anat., 1869, v, 49.
(10) TARCHANOFF: Beobachtungen über contractile elemente in den Blut und Lymphcapillaren. Pflüger's Arch., 1874, ix, 407.
(11) SEVERINI: Recerche sulla innervazione dei vasi sanguini. Perugia, 1878.
(12) MAYER: Ueber die blutleeren Gefässe im Schwanze der Batrachierlaven. Sitzungsb. d. math. naturwis. Classe d. Kaiserl. Akad. d. Wiss. in Wien, 1885, sci. abth. iii, 204.
(13) STEINACH AND KAHN: Echte Contractilität und Motorische Innervation der Blutcapillaren. Pflüger's Arch., 1903, xcvii, 105.
(14) ROY AND BROWN: The blood pressure and its variations in the arterioles, capillaries and smaller veins. Journ. Physiol., 1879, ii, 323.
(15) DALE AND RICHARDS: The vasodilator action of histamine and of some other substances. Journ. Physiol., 1918, lii, 110.
DALE AND LAIDLAW: Histamine shock. Journ. Physiol., 1918, lii, 355.
DALE, LAIDLAW AND RICHARDS: Traumatic toxaemia as a factor in shock. Med. Research Com., Special Rept. Series no. 26, 8.
(16) KROGH: The number and distribution of the capillaries in muscles, etc. Journ. Physiol., 1919, lii, 409.
KROGH: The supply of oxygen to the tissues and the regulation of the capillary circulation. Journ. Physiol., 1919, lii, 457.
KROGH: Studies on the capillarimotor mechanism. I. The reaction to stimuli and the innervation of the blood vessels in the tongue of the frog. Journ. Physiol., 1920, liii, 399.
(17) HOOKER: The functional activity of the capillaries and venules. Amer. Journ. Physiol., 1920, liv, 30.
(18) CLARK: Observations on living growing lymphatics in the tail of the frog larva. Anat. Record, 1909, iii, 183.
(19) BIEDL: Quoted from STEINACH and KAHN.
(20) FLEISCH: Zusammenfassende Betrachtungen über die Frage nach der Existenz einer aktiven Forderung des Blutstromes durch die Arterien. Schweizerische Med. Wochenschr., 1920, June 10, 323.
(21) HESS: Untersuchungen über den Antrieb des Blutstromes durch aktiven Gefässpulsationen. Pflüger's Arch., 1919, clxxiii, 243.
(22) DANZER AND HOOKER: Determination of the capillary blood pressure in man with the microcapillary tonometer. Amer. Journ. Physiol., 1920, lii, 136.
(23) CONNET: The effect of adrenalin on the venous blood pressure. Amer. Journ. Physiol., 1920, liv, 96.
(24) GUNN AND CHAVASSE: The action of adrenalin on veins. Proc. Roy. Soc., 1913, lxxxvi, B, 192.

(25) LISTER: On the early stages of inflammation. Phil. Trans., 1858, clxviii, 645.
(26) ROY: See ROY and BROWN, loc. cit.
(27) MAREY: Quoted from La Circulation du Sang, Paris, 1881, 377.
(28) MÜLLER: Studien über den Dermatographismus und dessen diagnostische Bedeutung. Deutsch. Zeitschr. f. Nervenheilk., 1913, xlvii–xlviii, 413.
(29) RYAN: The quantitative measurement of general fatigue. Amer. Journ. Physiol., 1918, xlv, 537.
(30) COTTON, SLADE AND LEWIS: Observations upon dermatographism with special reference to the contractile power of capillaries. Heart, 1917, vi, 227.
(31) BRISCOE: Observations on venous and capillary pressures with special reference to "Raynaud Phenomena." Heart, 1918, vii, 35.
(32) HOOKER: Observations on the venous blood pressure in man. Amer. Journ. Physiol., 1914, xxxv, 73.
(33) BRUCE: Vaso-dilator axon-reflexes. Quart. Journ. Exper. Physiol., 1913, vi, 339.
(34) BARDY: Ueber Hemmung inflammatorischer Symptome. Skand. Arch. f. Physiol., 1915, xxxii, 198.
(35) HOOKER: The veno-pressor mechanism. Amer. Journ. Physiol., 1918, xlvi, 591.
(36) HEUBNER: Ueber Vergiftung der Blutcapillaren. Arch. f. Exper. Path. u. Pharm., 1907, lvi, 370.
(37) ABEL AND KUBOTA: On the presence of histamine in the hypophysis cerebri and other tissues of the body, etc. Journ. Pharm. Exper. Therap., 1919, xiii, 243.
(38) CANNON: A consideration of the nature of wound shock. Journ. Amer. Med. Assoc., 1918, lxx, 611.
(39) RICH: Personal communication. Article to appear shortly in Journ. Exper. Med.
(40) LOMBARD: The blood pressure in the arterioles, capillaries and small veins of the human skin. Amer. Journ. Physiol., 1912, xxix, 335.
(41) THAYSEN: Ugeskrift f. Laeger, 1920, lxxxii, 473. Abstract, Journ. Amer. Med. Assoc., 1920, lxxv, 70.
(42) LAMSON: The part played by the liver in the regulation of blood volume and red corpuscle concentration in acute physiological conditions. Journ. Pharm. Exper. Therap., 1920, xvi, 125.
(43) SNYDER AND HAVENS: Changes in the blood after muscular activity and during training. Amer. Journ. Physiol., 1915, xxxvi, 239.
(44) SOLLMANN AND PILCHER: Endemic reactions. Journ. Pharm. Exper. Therap., 1917, ix, 309; 1917, ix, 350; 1917, x, 147.

THE CARBON DIOXIDE CARRIERS OF THE BLOOD

DONALD D. VAN SLYKE

The Hospital of The Rockefeller Institute for Medical Research

The arterial blood of man normally contains about 50 volumes per cent of carbon dioxide, the venous about 55 to 60. A pure solution containing 50 volumes per cent of carbon dioxide would have a hydrogen ion concentration of 3.1×10^{-5}. Yet in the blood the hydrogen ion concentration is only $\frac{1}{1000}$ as great, and the difference between arterial and venous blood is so slight as to be barely within the limits of measurement by our most sensitive methods. The purpose of the present paper is to discuss the chemical mechanism by which blood is enabled to carry its allotted load of carbon dioxide, and nevertheless maintain its hydrogen ion concentration at such a low and remarkably constant level.

Briefly stated, the organism utilizes two properties of certain constituents to set the blood pH[1] at the physiological point and, to maintain it near that point despite variations in CO_2 content such as occur in the change from venous to arterial blood. One of these properties is the buffer action of the bicarbonate and the proteins in the plasma, and of bicarbonate, proteins, and phosphates in the cells. The other property is the change from a relatively weak acid to a relatively strong one which hemoglobin undergoes when it changes from the reduced to the oxidized form. We shall discuss these factors separately, and then attempt to outline the manner in which they combine

[1] In speaking of reaction or hydrogen ion concentration we shall in this paper express data in the terms of pH introduced by Soerensen (1912) and generally adopted because of their convenience. The pH at a given reaction is the negative logarithm of the normality of the hydrogen ion concentration. The hydrogen ion concentration of a 0.01 N HCl solution, for example, is approximately 0.01 N, since the dissociation of HCl into H^+ and Cl' is practically complete. 0.01 is 10^{-2}. The logarithm of 10^{-2} is of course -2. The pH of 0.01 N HCl solution is therefore 2. The pH of water is almost exactly 7, which therefore represents the neutral reaction. A pH greater than 7 indicates an alkaline reaction, less than 7 an acid reaction. The reaction of normal blood at pH 7.35 is therefore slightly alkaline.

to enable the organism to meet both ordinary and extraordinary demands upon it for transport of carbon dioxide, and still maintain almost absolute constancy of hydrogen ion concentration in the blood.

FORMS IN WHICH CARBON DIOXIDE IS CARRIED IN THE BLOOD

There are four forms which CO_2 may assume in a solution, viz., dissolved anhydrous carbon dioxide, carbonic acid, bicarbonate, and carbonate. Using B to indicate a monovalent base, such as Na or K, these forms are indicated by the molecular formulas, CO_2, H_2CO_3, $BHCO_3$, and B_2CO_3 respectively. Concerning the equilibrium between the first two forms, CO_2 and H_2CO_3, we are uncertain. We can, however, assume that, if all the CO_2 in pure water solution is not changed to H_2CO_3, at least a constant proportion of it is. The reaction for the change is

$$CO_2 + H_2O = H_2CO_3.$$

From the mass law the conditions for equilibrium are

$$\text{(Concentration of } CO_2\text{)} \times \text{(Concentration of } H_2O\text{)} = \text{Constant} \times \text{(Concentration of } H_2CO_3\text{)}.$$

The H_2O concentration is constant, however, so the above equation becomes:

$$\text{Concentration of } CO_2 = \text{Constant} \times \text{Concentration of } H_2CO_3.$$

Since the H_2CO_3, therefore, either represents all the CO_2 not bound by bases, or represents a constant portion of it, it is customary in the study of equilibrium to assume that all of the CO_2 is in the form H_2CO_3. All calculations of the acid dissociation constant of H_2CO_3 have been based on this assumption, and no errors are introduced if in the use of this constant the same assumption is consistently made. Throughout the present paper, therefore, we shall follow the custom of indicating all the dissolved CO_2 not bound by bases as H_2CO_3.

Of the other two forms, $BHCO_3$ and B_2CO_3, only the bicarbonate can exist in significant quantities in solutions containing as much H_2CO_3 as does the blood. The reaction $B_2CO_3 + H_2CO_3 = 2BHCO_3$ possesses such an equilibrium constant that with physiological concentrations of H_2CO_3 it goes with practical completeness from left to right. This was shown by Bohr (1909):

> In a 0.155 per cent sodium carbonate solution (about the average carbonate concentration of plasma) at 38°, and with the physiologically normal carbon dioxide tension of 45 mm., Bohr calculated that 99.5 per cent of the sodium carbonate

was in the form of bicarbonate, and confirmed the calculation experimentally within the limit of analytical error. Even at 12 mm. CO_2 tension, which is seldom if ever observed in life except in premortal coma, 98 per cent was in the form of bicarbonate. Consequently one can, for quantitative purposes, regard the bicarbonate of the plasma as synonymous with its entire reserve of alkali in excess of that neutralized by acids other than carbonic.

In considering CO_2 transport by the blood we may therefore limit ourselves to two forms, free CO_2, indicated as H_2CO_3, and bicarbonate, $BHCO_3$. If all the CO_2 of the blood were in the free form as H_2CO_3, the blood would, as shown above, be 1000 times more acid than it is, or can be with compatibility to life. If all the CO_2 were in the form of carbonates, on the other hand, the blood would be hundreds of times too alkaline. The maintenance of the reaction necessary for life requires a definite and constant balance between CO_2 as H_2CO_3 and CO_2 as $BHCO_3$, and it is by means of its mechanism for maintaining such a balance that the organism is able to transport in the blood the great amounts of CO_2 that are continuously carried to the lungs for excretion, and to do so with a barely measurable change in the blood pH.

THE BUFFERS OF THE BLOOD

The nature of buffers

Buffers are salts of either weak bases or weak acids, and are characterized by the ability to enable solutions in which they are present to receive additions of limited amounts of either acid or alkali with much less change of pH than would be caused by the same additions to water, or to a solution containing only salts, such as NaCl, that have no buffer power. All the important buffers of the blood, viz., the bicarbonates, phosphates, and the alkali salts of the proteins, are salts of weak acids. In each buffer, part is present as free acid, part as the salt of a strong base, and the pH of the blood is determined by the relative proportions of buffer salts and free buffer acids respectively.

The alkali salts which constitute the buffers of the blood are in effect reservoirs of alkali, a portion of which they give up to neutralize carbonic or any other acid that enters the blood. In this manner act the phosphates, and also the alkali salts of the plasma proteins and of the hemoglobin.

In regard to such buffers two general laws may be stated, the principles underlying both of which may be found in L. J. Henderson's monograph on the regulation of body neutrality (1909).

1. The hydrogen ion concentration of the buffer solution is proportional to the ratio $\frac{\text{free acid}}{\text{free salt}}$ or $\frac{HA}{BA}$. Examples of such ratios in the blood are $\frac{H_2CO_3}{BHCO_3}$, $\frac{BH_2PO_4}{B_2HPO_4}$, $\frac{HHbO}{BHbO}$, $\frac{HHb}{BHb}$. (B is used to indicate any monovalent base, such as Na or K, A to indicate the acid radicle, BHbO the alkali salt of oxyhemoglobin, HHbO the free oxyhemoglobin, BHb and HHb the salt of reduced hemoglobin and the free protein respectively.)

2. A given buffer is most efficient in maintaining constancy of pH (that is, in minimizing the *proportion* by which H^+ is changed by a given addition of acid or alkali) when the ratio $\frac{HA}{BA}$ approximates 1, and H^+ equals K, the dissociation constant of the free acid forming one of the buffer pair.

The theoretical demonstration of these two laws is the following:

1. *Relationship between pH and the ratio* $\frac{HA}{BA}$ (or $\frac{\textit{free buffer acid}}{\textit{buffer salt}}$). This relationship is derived as follows:

Since the reaction of electrolytic dissociation of an acid into H^+ and anion is $HA = H^+ + A'$, it follows from the law of mass action that, at equilibrium,

$$1)\quad K \times HA = H^+ \times A',$$

K being the dissociation constant of the acid. Hence

$$2)\quad H^+ = K \times \frac{HA}{A'}$$

But when the buffer mixture is the salt of a very weak acid plus some free acid, only an infinitesimal part of the anion, A′, originates from dissociation of the free acid (H_2CO_3 in the concentration present in the blood, for example, is dissociated into H^+ and HCO_3' only to the extent of about 1/10,000). Essentially all of the anion originates from dissociation of the salt, BA, into B^+ and A′. Most salts in 0.1 to 0.01 molecular concentration undergo such dissociation to the extent of 60 to 90 per cent of the amount present. If the degree of dissociation be represented by λ, the concentration of anions is $A' = \lambda BA$, and Equation 2 may be written

$$3)\quad H^+ = K \frac{HA}{\lambda BA}.$$

Since λ varies to a relatively slight extent over ranges of concentration within such limits as are found in blood constituents, one may state as a close approximation that $\frac{K}{\lambda} = K_1$ and

$$4)\ H^+ = K_1 \frac{HA}{BA}.$$

In terms of pH, since $pH = -\log H^+$, Equation 4 is

$$5)\ pH = -\log K_1 - \log \frac{HA}{BA}, \text{ or}$$

$$6)\ pH = pK_1 + \log \frac{BA}{HA}.$$

The form represented in Equation 6 was first introduced by Hasselbalch (1916).

Expressing $-\log K_1$ as pK_1 the value of pK_1 for B_2HPO_4: BH_2PO_4 mixtures was shown by Soerensen (1909) and by Clark and Lubs (1916) to be approximately 6.8. For $BHCO_3$: H_2CO_3 solutions in the concentrations normally found in blood plasma Hasselbalch (1916) found pK_1 at body temperature to have a value of 6.1.[2] His results indicated in fact that by determination of the ratio $BHCO_3 : H_2CO_3$ the pH of a blood sample could be estimated with Equation 6 as accurately as by the standard electrometric method.

2. *Maximum Efficiency of Buffer Action When* $\frac{BA}{HA} = 1$, *and* $H^+ = K$.

There is a slight inconsistency in the above heading, because when $\frac{BA}{HA} = 1$, H^+ is not quite equal to K, the dissociation constant of the acid, HA. For, from Equation 3, $H^+ = \frac{K}{\lambda}\frac{HA}{BA}$. Hence when $H^+ = K$, $\frac{HA}{BA} = \lambda$ instead of $\frac{HA}{BA} = 1$. However, λ is so nearly equal to 1 in the buffer salts at the dilution in which they occur in the blood, that the error in considering it as unity may, for the purposes of the present discussion, be neglected.

[2] The figure actually given by Hasselbalch is 6.4, because he used the equivalent concentration of H_2CO_3, which is twice the molecular, in calculating the $\frac{BHCO_3}{H_2CO_3}$ ratio. When used with the molecular ratio adopted by L. J. Henderson and, so far as we have observed, all other authors except Hasselbalch, the value of pK_1 must, therefore, be reduced by log 2, or 0.3, giving pK_1 the value 6.1.

The fact that a buffer mixture of a weak acid and its salt has its maximum efficiency, when $\frac{BA}{HA} = 1$, in diminishing changes in reaction caused by adding either base or acid, has its basis in the general proposition, that in the ratio $\frac{a}{1-a}$ a given change in a produces the least change in the ratio when $a = 0.5$ and the ratio is consequently unity. The relationship is exemplified graphically by a curve expressing as ordinates values of the ratio $\frac{BA}{HA}$, as abscissae values of pH. For bicarbonate at approximately normal blood plasma (0.03 *M*) concentration, the curve indicated by figure 1 is obtained. The curve is calculated from the approximate equation $pH = 6.1 + \log \frac{BHCO_3}{H_2CO_3}$, the, for the present purposes, insignificant variations in pK_1 from the value 6.1 being neglected. It is evident from inspection that the curve is steepest in the middle, when H_2CO_3 and $BHCO_3$ are equal (50 per cent of the CO_2 as $BHCO_3$). This fact means that at this point the addition of sufficient acid or alkali to change a given amount of the total CO_2 from $BHCO_3$ to H_2CO_3, or vice versa, causes less shift in pH than at points near either end of the curve, where the ratio $BHCO_3$: H_2CO_3 is much greater or smaller than 1. For example, changing the per cent of CO_2 as $BHCO_3$ from 50 to 60 alters the pH from 6.10 to 6.26, a change of 0.16; while changing the percentage from 85 to 95 raises the pH from 6.85 to 7.40, a change of 0.55. It is evident that at pH 7.40, the reaction of normal blood plasma, when the ratio $\frac{BHCO_3}{H_2CO_3} = \frac{20}{1}$, bicarbonate as a buffer is acting at far from its most efficient point.

The phosphates are more efficient at blood pH. When the ratio $\frac{B_2HPO_4}{BH_2PO_4} = 1$, the pH is 6.8, much nearer to the reaction of normal blood than the 6.1 pH of the $BHCO_3 : H_2CO_3$ pair when their ratio is unity. The phosphates are present in so small an amount in plasma that they play a quantitatively negligible rôle, but they are more important in the cells. Of all the organic and inorganic acids, other than proteins, that might conceivably be used as buffers by the organisms, L. J. Henderson (1909) estimates that carbonic and phosphoric most nearly approximate maximum efficiency at the blood reaction, despite the fact that the blood pH is not very near the point of maxi-

mum efficiency of either buffer. It is of importance, however, that this point is in both buffers at a lower pH than that normal for blood, so that if the blood pH falls, the phosphate and carbonate buffers oppose the change with an efficiency which increases as the change approaches the danger point. At pH 6.95, which appears to be the most acid point consistent with life in man, the phosphates have very nearly their maximum buffer efficiency.

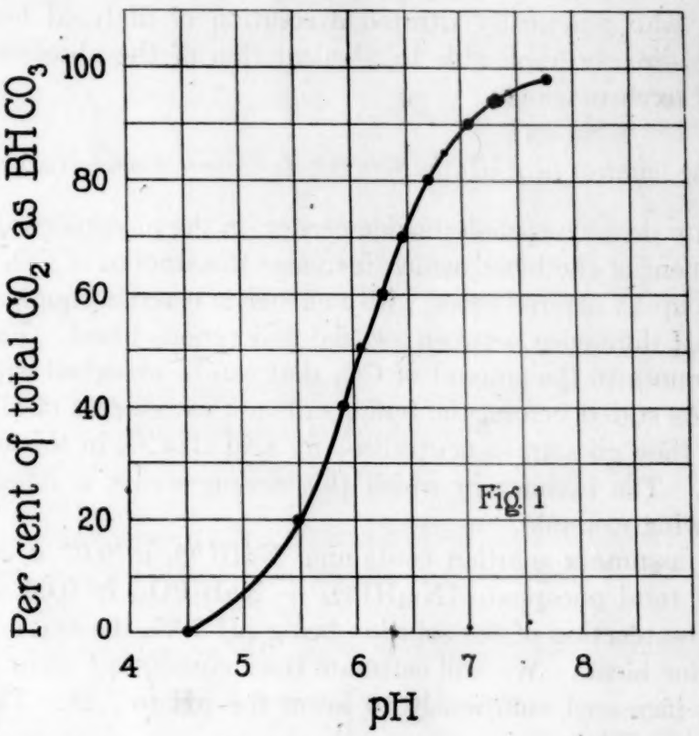

Fig. 1. Action of $NaHCO_3$:H_2CO_3 buffer, showing maximum buffer effect at middle of curve when $NaHCO_3$:H_2CO_3 ratio = 1.

While the organism has appropriated the best available, although not ideal, buffers that the external laboratories of nature offer, it appears to have manufactured, in the chief blood proteins, buffers of its own which are nearly ideal. For oxyhemoglobin Campbell and Poulton (1920) have recently estimated that the acid dissociation constant is approximately 2×10^{-8}, according to which the point of maximum buffer efficiency is at pH 7.7 which is close to that of blood. For reduced hemoglobin L. J. Henderson (1920) estimates that the acid dissociation constant is $\frac{1}{9}$ as great as for oxidized. This would

make it 2.2×10^{-9}, with a maximum buffer efficiency at pH 8.66, if we accept Campbell and Poulton's approximate figure for oxyhemoglobin. Henderson's own calculation placed the absolute values of both constants approximately 10 times as high. While his data, estimated from the differences between the CO_2 absorption curves of oxygenated and reduced blood found by Christiansen, Douglas and Haldane, seem adequate for approximating the difference in the K's of the two forms of hemoglobin, the experiments of Campbell and Poulton, who practically titrated a solution of dialyzed hemoglobin, are more directly applicable to a calculation of the absolute value of the K of oxyhemoglobin.

The manner in which buffers act as carbon dioxide carriers

We may define a carbon dioxide carrier, in the physiological sense, as a constituent of the blood which increases the amount of CO_2 that may be taken up by arterial blood with a change in reaction equal only to the normal pH difference between arterial and venous blood. Each buffer adds its quota to the amount of CO_2 that can be absorbed within these limits. As stated before, the buffer salts act as reserves of alkali, part of which they give up to neutralize any acid (H_2CO_3 in this case) that is added. The manner in which this action occurs is illustrated by the following example.

Let us assume a solution containing $NaHCO_3$ in 0.03 M concentration, and total phosphate ($Na_2HPO_4 + NaH_2PO_4$) in 0.05 M concentration, the reaction of the solution being pH 7.35, the average normal reaction for blood. We will calculate the changes that occur when the H_2CO_3 is increased sufficiently to lower the pH to 7.25. The molecular ratios $\frac{BA}{HA}$, at this reaction are calculated by rearranging the general equation, $pH = pK_1 + \log \frac{BA}{HA}$, to the form $\log \frac{BA}{HA} = pH - pK_1$. Using for pK_1 in the bicarbonate and phosphate buffers respectively the values 6.10 and 6.80 we have

$$\log \frac{Na_2HPO_4}{NaH_2PO_4} = 7.35 - 6.80 = 0.55$$

$$\frac{Na_2HPO_4}{NaH_2PO_4} = 3.55$$

The total concentration of PO_4 is 0.05 M. Therefore $NaH_2PO_4 = 0.05 - Na_2HPO_4$, whence

$$\frac{Na_2HPO_4}{0.05 - Na_2HPO_4} = 3.55$$

$$\left.\begin{aligned} Na_2HPO_4 &= 0.0390 \text{ M concentration} \\ NaH_2PO_4 &= 0.0110 \text{ M concentration} \end{aligned}\right\} \text{at pH } 7.35$$

In the same way we calculate

$$\left.\begin{aligned} Na_2HPO_4 &= 0.0346 \text{ M concentration} \\ NaH_2PO_4 &= 0.0154 \text{ M concentration} \end{aligned}\right\} \text{at pH} = 7.25$$

Since NaH_2PO_4, however, has 1 less Na than Na_2HPO_4, the change from the latter to the former sets free 1 equivalent of Na to combine with H_2CO_3 and form $NaHCO_3$. The concentration of alkali thus freed from the buffer reservoir is calculated as

$$\begin{aligned} &0.0390 \text{ M } Na_2HPO_4 \text{ at pH } 7.35 \\ &0.0346 \text{ M } Na_2HPO_4 \text{ at pH } 7.25 \\ \text{Difference} = &0.0044 \text{ M Na set free to form } NaHCO_3. \end{aligned}$$

At the beginning, with pH 7.35 our $NaHCO_3$ was 0.030 M. We have now by alkali drawn from the phosphates increased it by 0.0044 to 0.0344 M.

We have as additional CO_2 increase added a small amount of free H_2CO_3.

The $\frac{NaHCO_3}{H_2CO_3}$ ratio at the start is calculated as follows:

$$\log \frac{NaHCO_3}{H_2CO_3} = \text{pH} - 6.1 = 7.35 - 6.10 = 1.25$$

$$\frac{NaHCO_3}{H_2CO_3} = 17.8$$

$$NaHCO_3 = 0.03$$

$$H_2CO_3 = \frac{0.03}{17.8} = 0.00169 \text{ M}$$

At pH 7.25, the ratio $\frac{NaHCO_3}{H_2CO_3}$ is 14.1; hence if the $NaHCO_3$ had remained at 0.03 M, the H_2CO_3 would have been

$$H_2CO_3 = \frac{0.03}{14.1} = 0.00212 \text{ M}.$$

Since the $NaHCO_3$ is increased to 0.0344, however, we have at pH 7.25

$$H_2CO_3 = \frac{0.0344}{14.1} = 0.00244 \text{ M}.$$

The increase from 0.00169 to 0.00212 = 0.00043 M represents the amount of free H_2CO_3 that is added to 0.03 M $NaHCO_3$ solution in changing the pH from 7.35 to 7.25. It is the CO_2 carrying capacity of the bicarbonate alone within this pH range.

Due to the phosphate we have been able to add in addition

$$\begin{aligned} 0.00244 - 0.00212 = {} & 0.00440 \text{ M } CO_2 \text{ as } NaHCO_3, \text{ and} \\ & \underline{0.00032 \text{ M } CO_2 \text{ as } H_2CO_3} \\ & 0.00472 \text{ M } CO_2 = CO_2 \text{ capacity of the phosphate} \end{aligned}$$

between pH 7.35 and 7.25.

The results are summarized in table 1. The molecular concentrations are transformed into terms of volumes per cent of CO_2 by multiplying by 2240, since 100 cc. of a molecular solution contain 2240 cc. of CO_2 gas.

TABLE 1

	MOLECULAR CONCENTRATION × 1000		VOLUMES PER CENT CO_2		MILLIMETERS CO_2 TENSION CORRESPONDING TO H_2CO_3
	CO_2 as $NaHCO_3$	CO_2 as H_2CO_3	CO_2 as $NaHCO_3$	CO_2 as H_2CO_3	
CO_2 at pH 7.35	30.00	1.69	67.20	3.79	51.9
Increase on changing to pH 7.25					
Due to $NaHCO_3$	0.00	0.43	0.00	0.95	13.0
Due to phosphates	4.40	0.32	9.86	0.72	9.9
Total increase	4.40	0.75	9.86	1.67	22.9
CO_2 at pH 7.25	34.40	2.44	77.06	5.46	74.8

We have given this example in such detail because it illustrates exactly the manner in which the buffers of the blood act as CO_2 carriers.

Furthermore the bicarbonate concentration in the above is about that of blood plasma, the pH range is normal for blood, and the amount of phosphate present approximates in its CO_2 carrying effect the combined buffers of the blood.

Graphic measurement of carbon dioxide carrying capacity of buffers

Both the demonstration and measurement of the CO_2 carrying capacity of a buffer solution are facilitated by the use of a graphic method, which we shall later use to demonstrate the CO_2 carriers of the blood, and for the illustration of which the above data will serve.

If we draw a curve, expressing $BHCO_3$ values as ordinates, H_2CO_3 values as abscissae, the curve will be a straight line for all points corresponding to any given $\frac{BHCO_3}{H_2CO_3}$ ratio, and the line will rise more or less steeply according as the $\frac{BHCO_3}{H_2CO_3}$ ratio is great or small. But a constant $\frac{BHCO_3}{H_2CO_3}$ ratio indicates a constant pH. Consequently we are able by a series of straight, slanting lines on a diagram arranged as described to express all possible $\frac{BHCO_3}{H_2CO_3}$ ratios and pH values. Diagrams of this nature have been introduced by Haggard and Henderson (1919) for construction of their "CO_2 diagrams" of the blood.

For our purpose, we wish to express as ordinates the total $H_2CO_3 + BHCO_3$ rather than the $BHCO_3$ alone. The rectilinear pH line may still be used, but the calculation of the locus of the line is slightly more complicated. Since it is customary in physiological experiments to express the free CO_2 in terms of CO_2 tension, or the mm. of total atmospheric pressure due to the fraction which consists of CO_2, it is convenient also to use as abscissae units of CO_2 tension. It is permissible to use CO_2 tension in place of H_2CO_3 concentration because, according to the laws of gas solubility, the tension in the gas phase is proportional to the concentration in the solution.

Since at 38° (Bohr, 1905) water in contact with pure CO_2 at 760 mm. tension dissolves 55.5 volume per cent of CO_2, the free CO_2 (or H_2CO_3) of a water solution in equilibrium with CO_2 at 38° and p_{CO_2} mm. of tension may be calculated as:

$$\text{For water, volumes per cent } CO_2 \text{ as } H_2CO_3 = 55.5 \times \frac{p_{CO_2}}{760} = 0.073\ p_{CO_2}.$$

For blood serum Bohr found that the solubility coefficient was 0.540, slightly less than for water, and for whole blood it was 0.511. Consequently:

For blood serum, volumes per cent CO_2 as $H_2CO_3 = 54.0 \times \frac{p_{CO_2}}{760} = 0.0711\ p_{CO_2}$.

For whole blood, volumes per cent CO_2 as $H_2CO_3 = 51.1 \times \frac{p_{CO_2}}{760} = 0.0672\ p_{CO_2}$.

The total CO_2 content of a water solution under p_{CO_2} mm. of CO_2 tension is therefore calculated as $BHCO_3 + H_2CO_3 = BHCO_3 + 0.073\ p_{CO_2}$.

The ratio $\frac{BHCO_3}{H_2CO_3} = \frac{\text{Total } CO_2 - 0.073\ p_{CO_2}}{0.073\ p_{CO_2}}$.

Designating the ratio $\frac{BHCO_3}{H_2CO_3}$ as R we have

$$p_{CO_2} = \frac{\text{Total } CO_2}{0.073\ (R + 1)}.$$

For locating the pH lines, one calculates R for the desired pH values from the equation $\log R = pH - 6.10$, and then, inserting a convenient value for the total CO_2 in the above equation, calculates the tension, p_{CO_2}, corresponding to the point on the line through which the pH line passes.

If the chart is prepared for whole blood the factor 0.0672 is used in place of 0.073; if for serum or plasma, the factor 0.0711.

The calculation is really simple. For example, if we choose 75 as the total CO_2 line on which to lay off our tension values, the latter are calculated as

$$p_{CO_2} = \frac{75}{0.073\ (R + 1)} = \frac{1027}{R + 1}, \text{ for a water solution, or as}$$

$$p_{CO} = \frac{75}{0.0672\ (R + 1)} = \frac{1115}{R + 1} \text{ for whole blood.}$$

Figure 2 is drawn in the manner described. The level of point *A* represents the CO_2 content of the solution ($BHCO_3 + H_2CO_3 = 67.20 + 3.79 = 70.89$) at pH 7.35, *B* represents the content after the tension has been raised sufficiently to change the pH to 7.25. The vertical distance *A'B* represents the total capacity of the solution to take up CO_2 between the pH limits 7.35 and 7.25. The slight distance from the line *AA'* to *C* represents the CO_2 capacity of the bicarbonate alone, the remainder of the capacity is due to the phosphate.

The slant of the line *AC* represents the increase in free H_2CO_3 caused by increasing the CO_2 tension, the increase being 0.073 volume per cent for each mm. increase in CO_2 tension.

The vertical distance of point *C* above the line *AA'* represents the increase in H_2CO_3 sufficient to have brought the bicarbonate solution

to pH 7.25 had no buffer other than the bicarbonate been present; it therefore represents the relatively slight CO_2 carrying capacity of 0.03 M $NaHCO_3$ solution from pH 7.35 to 7.25.

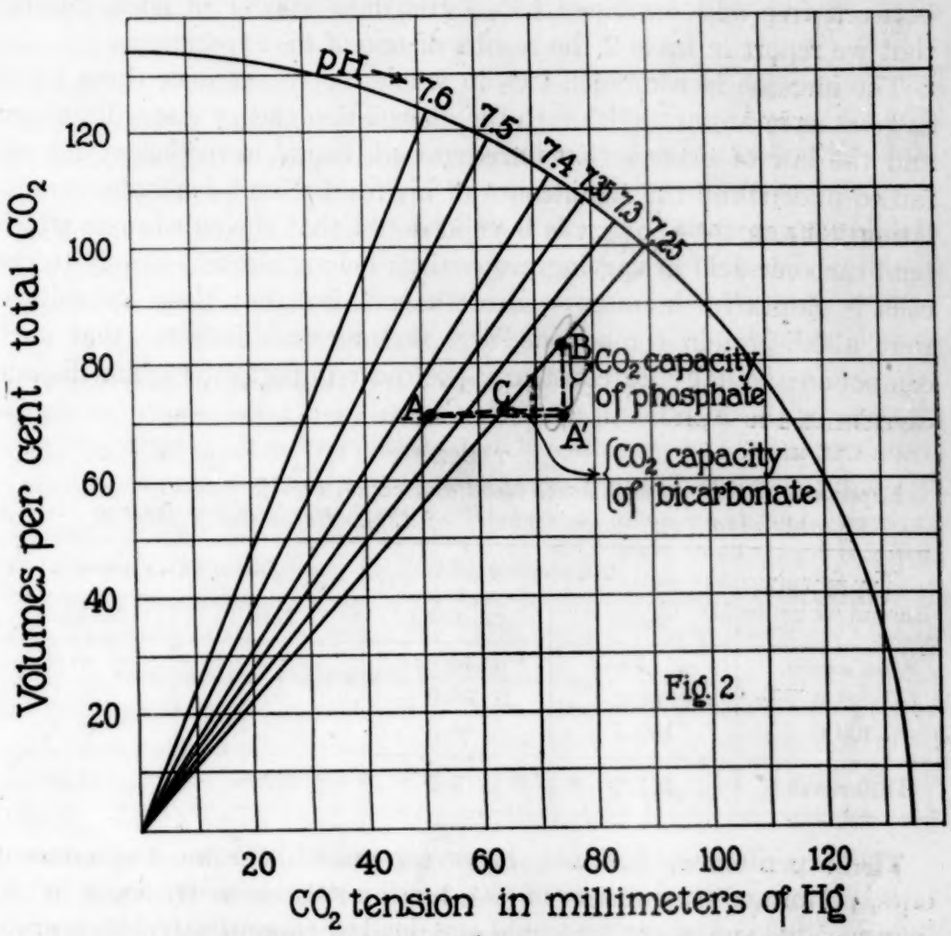

Fig. 2. CO_2 carrying power between pH 7.35 and pH 7.25 (indicated on chart as CO_2 capacity) of a solution containing 0.05 M phosphate and 0.03 M $NaHCO_3$.
Vertical distance from line A–A_1 to point B = total CO_2 power.
Vertical distance from line A–A_1 to point C = CO_2 carrying power of $NaHCO_3$.
Vertical distance from line C to B = CO_2 carrying power of phosphate.

The carbon dioxide carrying capacity of the buffers of the blood

Experimental evidence that both cells and plasma contain buffers other than bicarbonate and that the cells are several times richer in such buffers than the plasma, was published as long ago as 1868 by Zuntz. He determined the CO_2 absorption curves of whole blood and

separated serum, precisely as have Joffe and Poulton (1920), subtracted the physically dissolved CO_2 from the total CO_2, as we have done in this paper in plotting figure 3, and thereby differentiated between free and combined CO_2. His data appear of such interest that we report in table 2 the results of one of his experiments.

The increase in combined CO_2 in whole blood was over three times as great as in serum. Although the dissociation theory was still unborn and the law of mass action unrecognized, Zuntz nevertheless did not fail to understand the significance of his results, and concludes (translation) "From the above it is beyond doubt that the substances which bind carbonic acid in varying proportions belong almost entirely to the cells." Zuntz furthermore was of the opinion that these substances were alkali-protein compounds, and present data indicate that such compounds actually do constitute the most important carbon dioxide carriers in the blood.

TABLE 2

Experiment of Zuntz with horse blood and serum equilibrated at 0° with air containing varying percentages of CO_2. Barometer = 758–760

CO_2 IN AIR	CO_2 IN SERUM		CO_2 IN WHOLE BLOOD	
	Total	Combined	Total	Combined
vol. per cent	*vol. per cent*	*vol. per cent*	*vol. per cent*	*vol. per cent*
22.9	69.3	38.0	95.6	26.3
100.0	180.5	44.3	233.4	47.9
Difference....	111.2	**6.3**	137.8	**21.6**

The experiments of Zuntz, however, were like some more recent ones, performed over ranges of CO_2 tension such as never occur in the living body, and therefore could not lead to quantitatively accurate conclusions concerning the part which the "CO_2 binding substances" play in the actual physiological regulation of the blood reaction.

Joffe and Poulton (1920) have recently published experiments which are of a nature peculiarly well adapted to show the effectiveness of the blood buffers as carbon dioxide carriers at reaction ranges within physiological limits. Furthermore their results demonstrate another and most interesting fact, viz., that the richly buffered cells are able, under the influence of changes in reaction, to transfer their buffer effect to the relatively poorly buffered plasma, so that the latter, when in contact with the cells becomes as efficient a CO_2 carrier as the former. The cell buffers themselves are chiefly hemoglobin and phosphate.

The former cannot diffuse out into the plasma, and the latter does not. There must consequently be some mechanism by which the cell buffers can, so to speak, act at a distance. The mechanism by which this is attained is so important that we shall devote a later section to its discussion.

In figure 3 we have reproduced the CO_2 absorption curves published by Joffe and Poulton, together with pH lines arranged as in figure 2 to make possible a comparison of the effectiveness of the buffers of the three fluids acting over a definite reaction range, pH 7.35 to 7.25.

There are other reliable blood CO_2 absorption curves in the literature from which the CO_2 carrying capacity may be estimated, and the power of the cell buffers to affect the plasma has also been studied. The peculiar value of the data of Joffe and Poulton for our demonstration, however, lies in the fact that they determined on the same blood: *a*, the amounts of CO_2 absorbed by whole blood under varying CO_2 tensions; *b*, the CO_2 contents of serum samples centrifuged from the blood *after* the latter had been equilibrated at the varying CO_2 tensions (the curve so obtained the authors term the CO_2 absorption curve of "true serum"); *c*, the amounts of CO_2 absorbed by serum centrifugated from the cells of the freshly drawn blood, and then equilibrated, *separately* from the cells, with different tensions of CO_2. This is called the "separated serum."

In constructing figure 3 it has been necessary to draw two pairs of pH lines, one pair for whole blood, the other for serum, because of the differences in solubilities of CO_2 in whole blood and serum, respectively, which have been previously discussed in connection with the construction of such diagrams.

A glance at the slopes of the three curves in figure 3 is sufficient to show that the separated serum has only a fraction of the buffer effect, and consequent CO_2 carrying capacity, of the whole blood, while the true serum approximates the whole blood in these respects. The quantitative relationships of the three carbon dioxide carrying capacities are indicated graphically by the vertical distances, indicated by brackets, which show the volumes per cent of CO_2 taken up by whole blood, true serum, and separated serum respectively in passing from pH 7.35 to 7.25. This pH difference is greater than that ordinarily found between venous and arterial blood, but it does not appear to be outside the range of physiological possibilities, and we have preferred it to a shorter range for the present chart, in order that the differences may be large enough to keep the errors of calculation low.

In figure 3 the points SS_1, SS_2, SS_3, etc., have the same relative significance as the points A, C, and B respectively in figure 2. The vertical distance between the horizontal lines on which SS_1 and SS_3 are located indicates the total CO_2 carrying capacity of the separated serum within the pH limits 7.35 and 7.25; the vertical distance between the levels of SS_1 and SS_2 indicates the portion of the CO_2 carrying

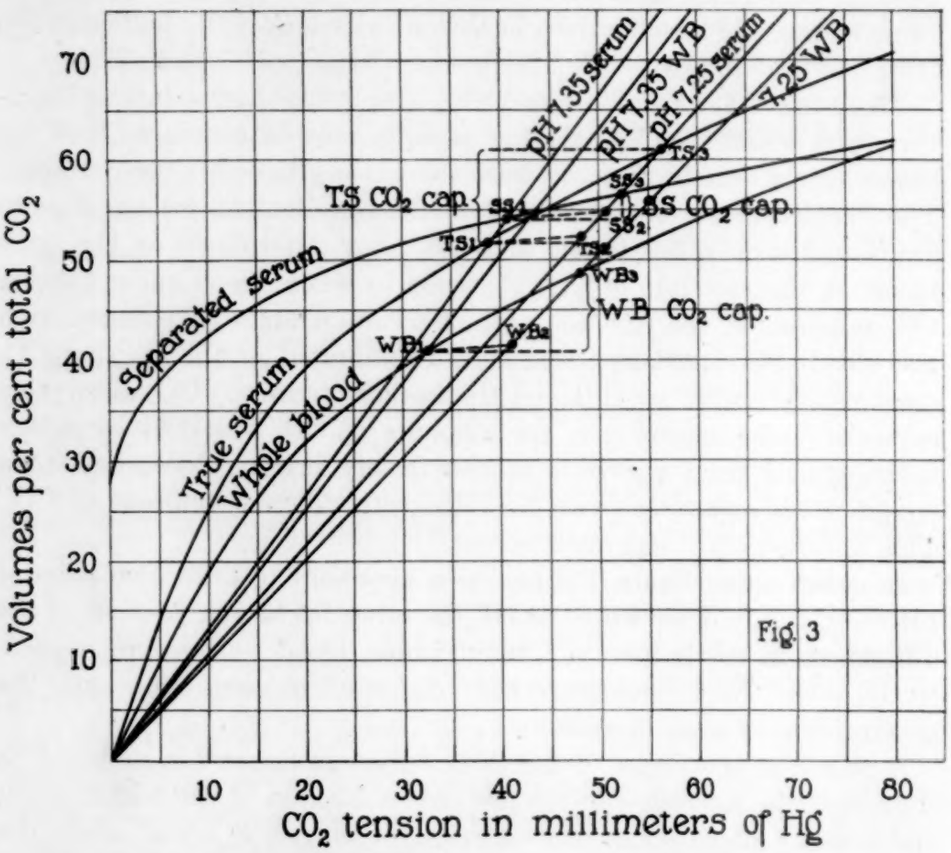

Fig. 3. CO_2 carrying powers (indicated as "CO_2 cap.") between pH 7.35 and pH 7.25 of oxygenated whole blood, of its true serum, and of separated serum. Data of Joffe and Poulton.

capacity that is due to the bicarbonate alone. Similar points on the true serum curve are indicated as TS_1, TS_2 and TS_3, respectively, and on the whole blood curve by WB_1, WB_2, and WB_3.

From the data carried from figure 3 to table 3 it is seen that the total CO_2 carrying capacities between pH 7.35 and 7.25 are, for whole blood, 8.0 volumes per cent, for "true serum" 9.4, and for "separated

serum" 2.6. When we subtract the amount of CO_2 that the bicarbonate present alone would take up, the values, due to other buffers, are 7.5, 8.8 and 2.0 volumes per cent.

It is evident, first, that within physiological limits as well as under Zuntz's experimental conditions, serum is less than one-third as richly buffered with CO_2 carriers as is the whole blood; and second, that when the same serum is in contact with the cells during a change in reaction, the cells loan it sufficient of their buffer effect to eliminate the difference between them. As a matter of fact, our calculations indicate that in changing from pH 7.35 to 7.25 the true serum took up even more CO_2 (9.4 volumes per cent) than the whole blood (8.0 volumes per cent).

TABLE 3

Results determined from figure 3

	WHOLE BLOOD		TRUE SERUM		SEPARATED SERUM	
	CO_2 tension	CO_2 content	CO_2 tension	CO_2 content	CO_2 tension	CO_2 content
	mm.	*vol. per cent*	*mm.*	*vol. per cent*	*mm.*	*vol. per cent*
At 7.35	32.4	40.8	38.5	51.6	40.5	54.1
At 7.25	48.0	48.8	56.0	61.0	52.0	51.5
Difference	15.6	8.0	17.5	9.4	11.5	2.6
CO_2 carrying capacity due to bicarbonate		0.5		0.6		0.6
CO_2 carrying capacity due to buffers other than bicarbonate		7.5		8.8		2.0

The difference hardly exceeds the possible error, however, in the difference between two CO_2 determinations, and we can merely say that in the oxygenated whole blood the CO_2 carrying buffer power is about evenly divided between serum and cells, although in actual buffer *content* the serum is less than one-third as rich as the whole blood, and therefore presumably less than one-sixth as rich as the separated cells.

Blood has been submitted to fairly complete chemical analyses, and it is practically certain that it contains no substances in considerable amount of which we do not have at least sufficient knowledge to tell whether or not they can act as buffers, i.e., whether or not they are salts of weak acids or bases. An examination of the constituents reveals among those present in amounts sufficient to have significant effect, only the proteins, the bicarbonate, and the phosphate, which can be expected to act as CO_2 carrying buffers.

The bicarbonate, as indicated by table 3, is responsible for only about $\frac{0.5}{8.0}$ or 6 per cent of the total CO_2 carrying power of oxygenated whole blood between pH 7.35 and 7.25.

Of the rôle played by the phosphate we must rely on rough calculation based on unsatisfactory data. So far as we know, determinations of the inorganic phosphate in human blood are not available. Abderhalden (1898), however, published such figures for ox blood, and we use them in want of more direct data. His results indicate 1.30 gram of inorganic P_2O_5 per 1000 grams of corpuscles, equivalent to 0.018 M PO_4''' concentration in the cells. In the plasma the amount is only 0.001 M. To estimate the part which the phosphate may play in the total CO_2 carrying power of the blood, we will assume that the phosphate is divided evenly through cells and plasma, as its buffer effect presumably is, reducing the PO_4''' to 0.009 M, and will compare this amount with that which gives the effect indicated in figure 2. In the latter, it is shown that a phosphate of 0.05 M concentration is capable of taking up 11 volumes per cent of CO_2 in passing from pH 7.35 to 7.25. A phosphate solution of 0.009 M is only 18 per cent as concentrated, and can be expected to take up about 0.18 as much CO_2 over the same pH range. $0.18 \times 11 = 2.0$ volumes per cent. This would attribute to the phosphates about 25 per cent of the CO_2 carrying capacity of oxygenated blood. We are, however, entirely uncertain as to whether Abderhalden's ox blood figures bear any relation to those of the human blood with which Joffe and Poulton worked, and the above estimation can serve only to indicate the general magnitude of the part that the CO_2 carrying power of phosphates may play in that of the whole blood.

Concerning the part which hemoglobin plays, Campbell and Poulton (1920) have published results which enable us to make an approximate computation. These authors have determined the CO_2 absorption curve of a solution of hemoglobin (laked red cells) which was dialyzed until free from diffusible substances. To the solution thus obtained sodium bicarbonate was added in amount sufficient to make its concentration 0.041 N, which gives a total base concentration to be bound by CO_2 and proteins together of the same order of magnitude as in blood. In figure 4 we have reproduced the curve of the oxyhemoglobin solution so that the CO_2 carrying power between pH 7.35 and 7.25 can be estimated as in figure 2. Since the protein concentration of about 8 per cent approximates that of the plasma, we have used the solubility figures for CO_2 in plasma in drawing the pH lines.

From figure 4 we see that the total CO_2 carrying capacity of the hemoglobin solution between pH 7.35 and 7.25 is about 4 volumes per cent, of which 1.0 volume per cent is due to the bicarbonate, 3 volumes per cent therefore to the hemoglobin. Campbell and Poulton determined that the hemoglobin content of their solution was 52 per cent that of normal blood on the Haldane scale. Hence to compare the

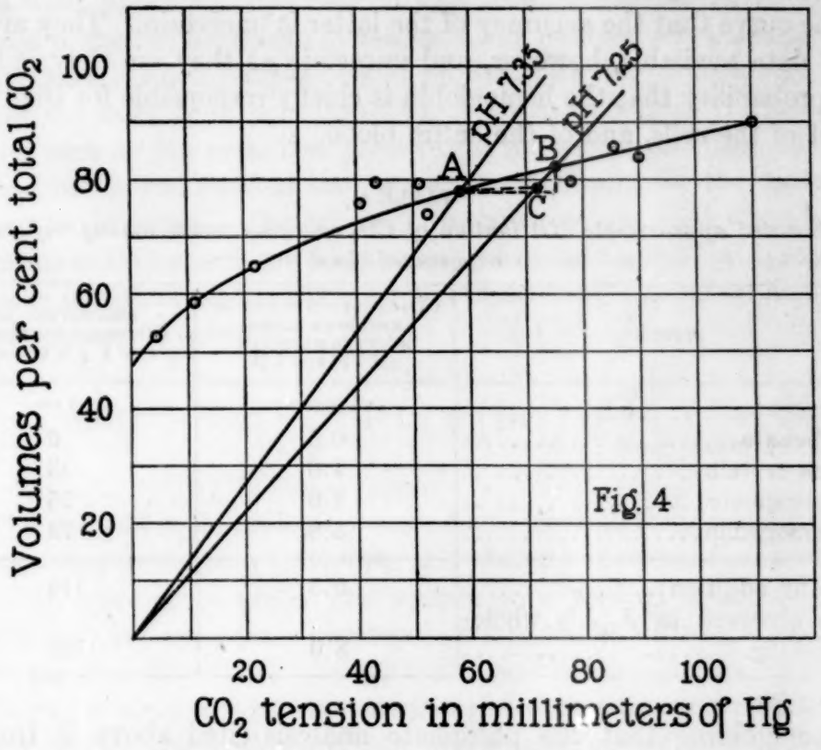

Fig. 4. CO_2 carrying power of a solution of dialyzed hemoglobin. Hemoglobin concentration corresponds to oxygen capacity of 9.6 volumes per cent. Data of Campbell and Poulton.

buffer effect with that of the hemoglobin in normal blood, we must multiply it by $\frac{1}{0.52}$, raising it to 5.8 volumes per cent. In the whole blood (table 3, fig. 3) we found 8.0 volumes per cent of CO_2 carrying power. If all the determinations and calculations were accurate, also the assumption that the blood used by Joffe and Poulton contained 100 per cent of hemoglobin, the above data would indicate that hemoglobin is responsible for $\frac{100 \times 5.8}{8.0} = 73$ per cent of the total CO_2 carrying power of the oxygenated blood.

We cannot state too emphatically that the data on which this calculation is based are only semi-quantitative. The exact hemoglobin content of the whole blood used by Joffe and Poulton, which we have used for comparison, is not given; it may have been 90 to 120 per cent of Haldane's average normal instead of 100 per cent, as is assumed. Furthermore, the individual CO_2 determinations of Campbell and Poulton on the dialyzed hemoglobin solution are so scattered on both sides of the curve that the accuracy of the latter is uncertain. They are the only data available, however, and uncertain as they are they indicate the probability that the hemoglobin is chiefly responsible for the buffer effect of the cells, and of the entire blood.

TABLE 4

Estimated approximate distribution of CO_2 carrying power among buffers of oxygenated blood

BUFFER	CO_2 CARRYING POWER OF SEPARATE BLOOD BUFFERS BETWEEN pH 7.35 and 7.25	PROPORTION OF CO_2 CARRYING POWER OF J. J.'S BLOOD
	vol. per cent	*per cent*
Bicarbonate	0.5	6
Plasma protein	1.0	13
Cell phosphate	2.0	25
Oxyhemoglobin	5.8	72
Total by addition	9.3	116
Total observed in J. J.'s whole blood (fig. 3)	8.0	100

Remembering that the phosphate analysis cited above is from a sample of ox blood, that the hemoglobin in the normal human blood was not determined and is only estimated to be the average normal, and that Campbell and Poulton's dissociation curve for dialyzed hemoglobin is not very accurate, it must be confessed that the data could hardly be less satisfactory for estimating the relative parts that the different buffers play in the total buffer effect of the blood. Taking the data for what they are, however, they indicate the distribution of the blood buffers shown in table 4.

DISCUSSION OF THE MECHANISM BY WHICH BUFFER EFFECTS ARE EXCHANGED BETWEEN CELLS AND PLASMA

The experiments of Zuntz (1868) cited above proved that the corpuscles are freely permeable to carbonic acid, but the first evidences that change in CO_2 tension causes a shift of other electrolytes between blood cells and plasma are apparently due to Gürber (1895). Gürber observed than when pure CO_2 is passed through blood the alkali titratable with methyl orange in the serum increases. The increase was presumably of bicarbonate. Such increase could be due either to diffusion of alkali out of the cells, or of acid into them. By ash analyses Gürber determined that no Na or K had passed into the serum, but that sufficient Cl had left it and entered the corpuscles to account for the increase in serum bicarbonate. The apparent explanation was that when minute amounts of HCl were formed by interaction of H_2CO_3 and NaCl, such HCl at once passed into the cells. Such a phenomenon could be illustrated by the following diagram:

$$H_2CO_3 + NaCl = NaHCO_3 + \left(\overline{HCl}\right)\Big|_{\text{wall}}^{\text{cell}} \Big|_{\text{interior}}^{\text{cell}} \longrightarrow$$

The effect of such a reaction would be to increase the plasma $NaHCO_3$ in response to increase in H_2CO_3, and thereby to cause the increase in the CO_2 carrying power of the plasma over that of separated serum which we have previously discussed in connection with Joffe and Poulton's results.

Gürber's experiments were conducted by passing pure CO_2 through blood, so that the range of CO_2 tension was altogether outside of physiological possibilities. Similar criticism applies to the recent experiments of Fridericia (1920). Van Slyke and Cullen (1917), however, found that the Cl transfer was sufficient to account for 72 per cent of the alkali increase in the plasma that was observed when the H_2CO_3 was raised from 0.0013 to 0.0024 M, both of the latter values being within the physiologically possible range. De Boer (1917) demonstrated that SO_4 anions are transferred in the same way as Cl; but the amount of sulfate in the plasma is so small, 0.002 M, that the proportion of it shifted could hardly have a measurable effect on the total alkali content. Hamburger (1916) has published evidence showing that to some extent cations, Na and K, take part in the shift. The cation shift must be at most, however, slight compared with the anion transfer. It may be

stated that the acid-base transfer between cells and plasma is carried out chiefly by the anions Cl' and HCO'_3.

Below is a diagram which although from lack of available data, it is necessarily incomplete, nevertheless appears to represent the chief buffer factors that maintain the constancy of the blood pH, and to indicate the anion exchange that makes the cell buffers available to the plasma. The equilibria as written are all shifted from left to right by increase in H_2CO_3.

PLASMA	CELL WALL	CELL
(1) $H_2CO_3 + NaCl \rightleftarrows NaHCO_3 + HCl$	$\rightarrow HCl \rightarrow$	(3) $HCl + K_2HPO_4 \rightleftarrows KH_2PO_4 + KCl$
(2) $H_2CO_3 + Na\ Protein \rightleftarrows NaHCO_3 + H\ Protein$		(4) $HCl + KHbO \rightleftarrows HHbO + KCl$
		(5) $HCl + KHb \rightleftarrows HHb + KCl$
	$\rightarrow H_2CO_3 \rightarrow$	(6) $H_2CO_3 + K_2HPO_4 \rightleftarrows KH_2PO_4 + KHCO_3$
		(7) $H_2CO_3 + KHbO \rightleftarrows HHbO + KHCO_3$
		(8) $H_2CO_3 + KHb \rightleftarrows HHb + KHCO_3$

"Na Protein" is used to indicate the total alkali salts of the plasma proteins, "H Protein" the free proteins not combined with base. Similarly "KHbO" and "HHbO" are used to indicate the alkali salt of oxyhemoglobin and the oxyhemoglobin not combined with alkali, respectively. "KHb" and "HHb" are used similarly for reduced hemoglobin. In the plasma the base is indicated as Na, in the cells as K, to indicate the fact that, in human blood at least, these two bases predominate in the two locations, and do not diffuse freely from one to the other. There is, of course, some K in the plasma, and some Na in the cells, which the above diagram fails to show.

In regard to the quantitative relationships among the reactions indicated in the diagram, our incomplete knowledge permits the following statements.

Reaction 1, with perhaps lesser reactions which end in other electrolyte transfers to the cells, accounts for two-thirds to three-fourths of the total bicarbonate formed in the plasma of whole blood by a rise in

H_2CO_3 such as to cause a pH change within the normal range. This is shown by the fact that an increase in H_2CO_3, sufficient to change the pH from 7.35 to 7.25, caused in the separated serum of Joffe and Poulton less than one-third as great a rise in $BHCO_3$ as in the true serum, in which the cells were present. The removal of the HCl from plasma by transfer into the cells makes it possible for Reaction 1 to proceed to a measurable degree. When cells are absent it cannot do so, and in the separated plasma only Reaction 2 can occur to a measurable extent.

Reaction 2, it follows, accounts for one-third to one-fourth of the $BHCO_3$ formed in the plasma of whole blood when the H_2CO_3 is increased within physiological limits. In the separated serum, presumably it accounts for practically the whole rise.

In the cells, the sum of the total $BHCO_3$ formed by Reactions 6, 7, and 8 is approximately equal to the $BHCO_3$ formed by Reactions 1 and 2 in the plasma (see for example parallel slopes of "true serum" and whole blood curves in figure 3). If the HCl shift is equivalent to $\frac{2}{3}$ to $\frac{3}{4}$ of the $NaHCO_3$ formed by Reactions 1 and 2, it follows, from the above paragraph, that Reactions 3, 4, and 5 together are equivalent to about $\frac{2}{3}$ to $\frac{3}{4}$ of Reactions 6, 7, and 8 combined. If this proves correct, then, for a given amount of bicarbonate formed in the cells, $\frac{2}{3}$ to $\frac{3}{4}$ as much chloride is taken in from the plasma.

The CO_2 carrying power of the phosphates in the cells, if PO_4 should approximate the same concentration in human as in ox cells (0.018 M) would be about $\frac{1}{3}$ that of the oxyhemoglobin for a change of pH over the range 7.35–7.25 (see table 4). Reactions 3 and 6 together would then be about equivalent to $\frac{1}{3}$ of the other 4 reactions in the cells. The data on the hemoglobin buffer effect and on the phosphate concentration are so uncertain, however, that the above estimate has only the validity of a rough approximation.

Furthermore, the data used have been obtained entirely with oxygenated hemoglobin, so that only Reactions 4 and 7 with hemoglobin were concerned. By reduction of the oxyhemoglobin, or part of it, so that Reactions 5 and 8 come into play, relationships are essentially altered, since, as previously stated (p. 147) reduced hemoglogin as a buffer is quite different from oxyhemoglobin. The effect of oxygen changes on the buffer action of hemoglobin is of great importance in regulating the blood reaction, and we shall consider it in the following section.

THE PART OF THE OXYGEN EXCHANGE IN THE CARBON DIOXIDE CARRYING CAPACITY OF THE BLOOD

We have shown that the oxygenated normal blood investigated by Joffe and Poulton could in changing its reaction from pH 7.35 to 7.25 take up 8 or 9 volumes per cent of CO_2. Half as much CO_2 would be about the normal difference between arterial and venous blood in a resting man, and would involve a pH change of only 0.05. The normal arterial-to-venous change was calculated by Parsons (1917) from electrometric measurements, however, to be only 0.02, and Peters and Barr (1921) in three normal men calculated differences of only 0.01 to 0.00. It appears that there is still another factor, besides those which we have studied, which assists in neutralizing the carbonic acid that the blood receives in the tissues.

Christiansen, Douglas and Haldane (1914) were led to a similar conclusion by a comparison of the CO_2 tension changes on the absorption curve of oxygenated blood with the differences estimated between arterial and venous CO_2 tensions in the lungs. The latter averaged 5 to 6 mm., while changes of 15 or 16 mm. were required to make oxygenated blood at 38° *in vitro* take up as much CO_2 as does the venous blood *in vivo*. The added factor in stabilizing the CO_2 tension was demonstrated in a notable paper by these authors to be the change which hemoglobin undergoes as it loses oxygen to the tissues and absorbs it in the lungs while carbon dioxide is passing in the reverse directions. These authors found that normal blood at 38° and any given CO_2 tension within the extreme physiological range (30–70 mm.) absorbed from 5 to 6 volumes per cent more CO_2 if the gas mixture with which the blood was equilibrated was $CO_2 + H_2$, than if the gas mixture was CO_2 + air. The only conceivable effect of the hydrogen as compared with the air was in its effect on the state of oxygenation of the hemoglobin. The conclusion was reached that the change from oxy- to reduced hemoglobin in the tissues enables the blood to take its required load of CO_2 with relatively as small an increase in CO_2 tension as that (5–6 mm.) which these authors calculated from alveolar air analyses.

From the data of Christiansen, Douglas, and Haldane, it is possible to demonstrate, as a matter of fact, that reduction of the blood can enable it to take up 5.5 volumes per cent of CO_2 and still not only maintain, but increase the alkalinity, because the increased CO_2 is all in the form of bicarbonate. At 40 mm. CO_2 tension, their curves give the total CO_2 contents tabulated below, from which the pH values are calculated as follows:

	TOTAL CO_2 CONTENT AT 40 MM. CO_2 TENSION	CO_2 AS H_2CO_3 = 40 × 0.0672	CO_2 AS $BHCO_2$	$pH = 6.10 + \log \frac{BHCO}{H_2CO_3}$
	vol. per cent	*vol. per cent*	*vol. per cent*	
Oxygenated blood..	51.5	2.69	48.8	7.36
Reduced blood.....	57.0	2.69	54.3	7.40

It is evident that H_2CO_3 must be added to the blood when it is reduced, or it will actually become more alkaline.

If we accept Campbell and Poulton's estimate of the acid dissociation constant of oxyhemoglobin, and accept L. J. Henderson's estimate of the smaller constant for reduced hemoglobin, we may calculate the amount of alkali which the hemoglobin of normal blood would give up, *a*, in passing from pH 7.35 to pH 7.33; *b*, in losing its oxygen and changing to reduced hemoglobin. Campbell and Poulton estimated from the curve we have reproduced in figure 4 that when all the hemoglobin in their solution was in the form BHbO, the amount of alkali thus bound was sufficient to combine as bicarbonate with 51.3 cc. of CO_2 per 100 cc. of solution. Since their solution contained only 52 per cent as much hemoglobin as normal blood, the alkali binding power of the hemoglobin in blood may be estimated as equivalent to $\frac{51.3}{0.52}$ = 99 cc. of CO_2 per 100 cc. of blood.

We may then estimate, by the method previously illustrated on pages 148–150 in calculating the CO_2 capacity of a bicarbonate-phosphate solution, the amount of base which the hemoglobin will yield to combine with H_2CO_3 in changing over a definite pH range. As such a range we shall take that from pH 7.35 to 7.33, since the normal difference between arterial and venous blood is found to be only pH 0.02 (Parsons, 1917). We shall express the data in terms of CO_2 equivalents.

The acid dissociation constants, K_O and K_R, of oxygenated and reduced hemoglobin respectively, as calculated by L. J. Henderson (1920) require revision in the light of the results of Campbell and Poulton (1920), who showed that the base-binding power of the oxyhemoglobin in 100 cc. of normal blood is equivalent to that of 99 cc. of CO_2, instead of the 18 cc. assumed, in the lack of other available data, by Henderson. Inserting 99 in place of 18 as a constant in Henderson's (1920) calculations, we have $K_O = 6.9 \times 10^{-8}$ and $K_R = 4.9 \times 10^{-8}$. The negative logarithms of these values are $pK_O = 7.16$ and $pK_R = 7.31$. Using these values, we calculate as follows:

$$\text{Log}\ \frac{\text{BHbO}}{\text{HHbO}} = 7.35 - 7.16 = 0.19$$

$$\frac{\text{BHbO}}{\text{HHbO}} = 1.55$$

$$\text{BHbO} = 60.2$$

Similarly at pH 7.33

$$\text{Log}\ \frac{\text{BHbO}}{\text{HHbO}} = 7.33 - 7.16 = 0.17, \text{ whence BHbO} = 59.1$$

If we now reduce all of the oxyhemoglobin to reduced hemoglobin, we have with pH still remaining at 7.33

$$\text{Log}\ \frac{\text{BHb}}{\text{HHb}} = 7.33 - 7.31 = 0.02, \text{ whence BHb} = 50.6$$

We have therefore yielded

1. By oxyhemoglobin in changing from pH 7.35 to 7.33 sufficient base to combine with 60.2 − 59.1 = 1.1 cc. of CO_2.

2. By oxyhemoglobin in being altered at pH 7.33 to reduced hemoglobin, sufficient base to combine with 59.1 − 50.6 = 8.5 cc. of CO_2.

Normally, however, only about one-third of the oxyhemoglobin is reduced in passing from arteries to veins, so we divide the 8.5 cc. by 3, making 2.83 cc. of CO_2 estimated as bound by the alkali set free as a result of the loss of oxygen.

We therefore estimate that the hemoglobin of normal blood, in passing from pH 7.35 to 7.33, and in being one-third reduced from oxy- to reduced hemoglobin, frees sufficient alkali to combine with 1.10 + 2.83 = 3.93 cc. of CO_2 per 100 cc. of blood. If we assign to the blood buffers other than oxyhemoglobin one-half of its CO_2 carrying capacity over the given pH change, as seems to be approximately correct, they will add 0.55 cc. to the above amount, making 4.48 cc. of CO_2 taken up. To this we must add the CO_2 taken up as H_2CO_3. At pH 7.33 the ratio $\frac{BHCO_3}{H_2CO_3}$ is $\frac{17}{1}$. Consequently in increasing the $BHCO_3$ by 4.48, we increase the H_2CO_3 by $\frac{4.48}{17}$ = 0.26 cc., making the total increase 4.48 + 0.26 = 4.74 cc. of CO_2 per 100 cc. of blood. The increase in

CO_2 tension accompanying the entire change is $\frac{0.26}{0.0672} = 3.9$ mm. The respiratory quotient, $\frac{CO_2}{O_2}$, is $\frac{4.74}{6.0} = 0.79$, a normal value.

The maintenance of the normal gas exchange of the organism with a difference of only 0.02 in pH between arterial and venous bloods is therefore explainable on the basis of the available physico-chemical data. The maintenance of an absolutely isohydric exchange, however, would not be thus explainable. It would require that without change in pH the absorption of each cc. of O_2 should cause the loss of at least 0.7 cc. of CO_2, since for each cc. of O_2 consumed the organism forms from 0.7 to 1.0 cc. of CO_2. The data of Christiansen, Douglas, and Haldane (1914) as shown in figure 2 of L. J. Henderson's paper (1920) indicate that at the normal H^+ value of 4×10^{-8} (pH = 7.4) complete loss of O_2 (18 cc.) causes isohydric absorption of only 8 cc. of CO_2, instead of the minimum required volume of $0.7 \times 18 = 12.6$ cc.

L. J. Henderson calculates that, although hemoglobin, like all proteins, is amphoteric and therefore capable of combining with acids, yet under the conditions of CO_2 tension and pH that exist *in vivo*, it cannot possibly combine directly with more than 1 or 2 volumes per cent of CO_2. As CO_2 carriers oxygenated and reduced hemoglobin function not through direct combination of the protein molecule with CO_2, but through acting as buffers, holding reserves of K and Na, which are given off to transform H_2CO_3 into $BHCO_3$ according to the laws of mass action whenever the H_2CO_3 concentration is increased. Further confirmation of this viewpoint is afforded by the results of Campbell and Joffe (1920) who showed that so long as conditions keep the pH within the normal physiological range, addition of oxyhemoglobin to alkali solutions does not increase, but diminishes, the amount of CO_2 combined at a given tension, the oxyhemoglobin competing with the CO_2 for the alkali, which the two substances divide according to their relative strengths as acids. It is only at pH ranges more acid than can exist in the circulation that hemoglobin combines with significant amounts of CO_2. We believe that data in the papers of Henderson, and of Campbell and Poulton, make untenable the view that hemoglobin acts to a significant extent as a CO_2 carrier by combining directly with the CO_2 in the blood.

Haggard and Y. Henderson (1920) have recently thrown some doubt on the validity of applying physiologically the results of Christiansen, Douglas and Haldane. Haggard and Henderson tend to the opinion that the peculiar effect of oxygenation and reduction on the CO_2 absorption curve of blood may be the result of a property artificially conferred on the blood by the process of defibrination, which was used in all of Christiansen, Douglas, and Haldane's experiments. In support of this view they give data indicating that oxalated blood, which they consider to be less altered from the natural circulating fluid than is defibrinated blood, does not show any oxygen effect on its CO_2 absorption curve. The assumption that oxalated blood is nearer to nature than defibrinated might readily be

questioned. The experimental fact that oxalated blood does not show the oxygen effect, however, is one which Peters has attempted to confirm in the writer's laboratory with consistently negative results: defibrinated and oxalated bloods have shown within the limit of experimental error identical oxygen effects on their absorption curves. It does not appear to the writer that any valid reason exists for doubting the physiological significance of the results of Christiansen, Douglas, and Haldane.

THE COMBINED EFFECT OF THE BUFFERS AND THE OXYGEN EXCHANGE IN ENABLING THE BLOOD TO CARRY ITS CARBON DIOXIDE WITH A NORMAL CONSTANCY OF REACTION

It remains to show the combined effect of the factors, which we have discussed in the preceding pages, in enabling the blood to take up the amounts of carbon dioxide which it does in passing from the arteries to the veins with the slight amount of reaction change which is observed.

For this demonstration we use the published CO_2 absorption curves of J. S. H. (Christiansen, Douglas and Haldane, 1914) and of J. J. (Joffe and Poulton, 1920). By means of them we shall estimate the changes which the blood undergoes when, starting completely oxygenated and with a reaction of pH 7.35, it progressively absorbs CO_2 and simultaneously loses volumes of oxygen sufficiently larger to maintain the normal respiratory quotient, 0.8.

In allowing for the effect, discussed in the preceding section, of the change from oxyhemoglobin to reduced hemoglobin on the CO_2 absorption curves, we have used an empirical formula which Peters (Peters and Barr, 1921) has devised. It expresses two assumptions, which hold with a close approximation to constancy over the normal range of CO_2 tension. (1) that the increase in volumes per cent total CO_2 caused by complete reduction of oxygenated blood at constant CO_2 tension is constant; (2) that for partial reductions, the effect on the CO_2 content is proportional to the reduction. In totally reduced blood the oxygen unsaturation is equal to the oxygen capacity, or 18.5 cc. of O_2 per 100 cc. of blood. (The oxygen unsaturation is the difference between oxygen content and total oxygen capacity, and represents the reduced hemoglobin.) If the increase in CO_2 content caused by total reduction of the hemoglobin is represented by N, the increase in CO_2 caused by an intermediate reduction is estimated as equal to the volume per cent of oxygen unsaturation $\times \frac{N}{18.5}$. From the data of Christiansen, Douglas and Haldane (1914) and of Joffe and Poulton (1920) Peters finds the average value of N for CO_2 tenslons between 30 and 70 mm. to be 6.5, and $\frac{N}{18.5} = 0.35$.

If we indicate by D the CO_2 increase due to oxygen loss,

$$D = 0.35 \times \text{volumes per cent of oxygen unsaturation.}$$

A theoretically exact formula could be based on the dissociation constants of the two forms of hemoglobin and the mass law principles expressed in the equations $pH = pK_1 + \log \frac{BHbO}{HHbO}$, and $pH = pK_1 + \log \frac{BHb}{HHb}$, but, considering the uncertainty of the exact value of pK_1 in at least the second of these equations, and the accuracy and simplicity of Peters' empirical formula, the latter seems at present preferable.

The relationship, between the increase in CO_2 content caused by any given increase in CO_2 tension in the oxygenated blood, and the increase in CO_2 content caused when the CO_2 increase is accompanied by an oxygen loss 1.25 times as great, is calculated as follows.

Let ${}_O\Delta_{CO_2}$ be the increase in CO_2 content of oxygenated blood caused by a given rise in CO_2 tension, without loss of oxygen, and ${}_R\Delta_{CO_2}$ be the increase in blood CO_2 content caused by the same rise in CO_2 tension, with the simultaneous loss of sufficient oxygen to give a respiratory quotient of 0.8.

O_2 = volumes per cent of oxygen lost, therefore of oxygen unsaturation.

Then $${}_R\Delta_{CO_2} = {}_O\Delta_{CO_2} + 0.35 \times O_2$$

$$= {}_O\Delta_{CO_2} + 0.35 \times \frac{{}_R\Delta_{CO_2}}{0.8}$$

Whence $${}_R\Delta_{CO_2} = 1.78\ {}_O\Delta_{CO_2}$$

Expressed in words the above equation means that the increase in total CO_2 of the blood, caused by a given rise in CO_2 tension, is further increased by 78 per cent if the blood loses during the change a volume of oxygen equal to 1.25 times that of the CO_2 it absorbs. In order to plot the actual CO_2 absorption curve of blood as it changes from arterial to venous, we must therefore add to each CO_2 increase measured on the absorption curve for oxygenated blood, 78 per cent more in order to allow for the effect of the normally accompanying oxygen loss.

In each chart (figs. 5 and 6) the heavy black curve AC represents the curve thus calculated, which may be called the true CO_2 absorption curve of blood. The CO_2 absorption curve of oxygenated blood is never touched in the body, save at point A by arterial blood. In the circulation the blood ordinarily travels for about one-third of the length of the true absorption curve from A toward C, and then returns (in the lungs) to its starting point at or near A. Under conditions of stress (e.g., exer-

cise) the venous blood moves further toward C, but apparently never, except in anemic individuals, travels more than about three-fourths of the distance AC.

In figures 5 and 6 we have drawn the pH lines through three points on the true CO_2 absorption curve AC. One is at A, which represents arterial blood. A is arbitrarily located in both bloods, for the purpose of

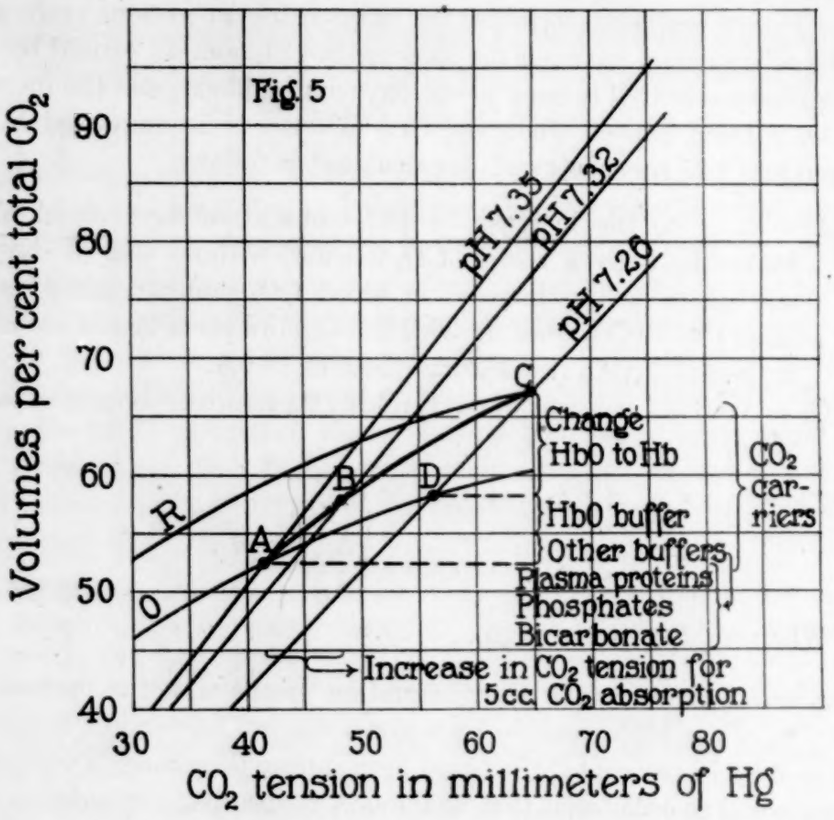

Fig. 5. Relative amounts of CO_2 carried by different agents in blood of J. S. H. The CO_2 absorption curve of oxygenated blood is marked O, that of wholly reduced blood is marked R.

Curve $A\ C$ represents estimated absorption curve of blood in the body, when the respiratory quotient is 0.8. Entire curve $A\ C$ represents maximum gas exchange, all of oxygen (18.5 cc.) being given off, and 14.8 cc. of CO_2 absorbed, with a pH change from 7.35 to 7.26.

Curve $A\ B$ represents absorption of 5 cc. of CO_2, with change of pH from 7.35 to 7.32.

$A\ D$ represents absorption of CO_2 that could occur in change from pH 7.35 to 7.26 if no hemoglobin were reduced.

The brackets on the right indicate the relative amounts of the added CO_2 neutralized at pH 7.26 by alkali supplied by the different carriers.

this demonstration, at pH 7.35, the average reaction of normal blood. A second point, *B*, indicates the conditions attending the absorption of 5 cc. of CO_2, and the loss of 6.3 cc. of oxygen, about the amounts that are ordinarily exchanged during a circulation. The third point, *C*, indicates the limit of the possible gas exchange, as at this point the

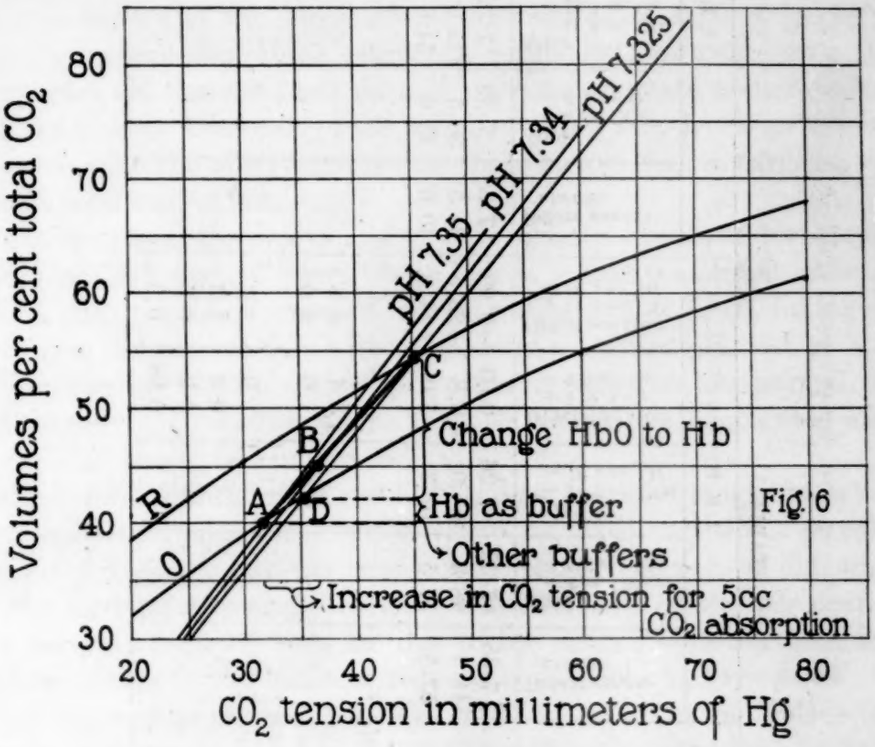

Fig. 6. Relative amounts of CO_2 carried by different agents in the blood of J. J.

The CO_2 absorption curve of oxygenated blood is marked *O*, that of wholly reduced blood is marked *R*.

Curve *A C* represents estimated CO_2 absorption curve of blood in the body, when the respiratory quotient is 0.8. Entire curve *A C* represents maximum gas exchange, all of the O_2 (18.5 cc.) being given off, and 0.8 as much, or 14.8 cc. absorbed, with a pH change from 7.35 to 7.325.

Curve *A B* represents absorption of 5 cc. of CO_2, with loss of 6.3 cc. of O_2, increase of 5.0 mm. in CO_2 tension, and change of pH from 7.35 to 7.34.

A D represents absorption of CO_2 that could occur in change from pH 7.35 to 7.325 if no hemoglobin were reduced.

If the slope of the CO_2 absorption curves were slightly greater, the line *A C* in the blood would fall upon the line of pH 7.35, and the gas exchange would be accomplished without change in pH. In this case, the HbO $\rightarrow$Hb change would furnish all the alkali, the buffers none.

TABLE 5

	J. B. H.					J. J.				
	Before gas exchange	After + 5 cc. CO_2 − 6.3 cc. O_2	After + 14.8 cc. CO_2 − 18.5 cc. O_2	Difference after + 5 cc. CO_2 − 6.3 cc. O_2	Difference after + 14.8 cc. CO_2 − 18.5 cc. O_2	Before gas exchange	After + 5 cc. CO_2 − 6.3 cc. O_2	After + 14.8 cc. CO_2 − 18.5 cc. O_2	Difference after + 5 cc. CO_2 − 6.3 cc. O_2	Difference after + 14.8 cc. CO_2 − 18.5 cc. O_2
pH	7.35	7.32	7.26	−0.03	−0.09	7.35	7.39	7.325	−0.01	−0.025
CO_2 tension, *mm*	41.5	48.3	65.0	+6.8	+23.5	31.5	36.5	45.0	+5.0	+13.5
Total CO_2, *vol. per cent.*	52.5	57.5	67.3	+5.0	+14.8	40.0	45.0	54.8	+5.0	+14.8
CO_2 increase due to HbO → Hb reaction, *vol. per cent*				3.4	8.8				3.8	12.6
CO_2 increase due to buffers, *vol. per cent*				1.6	6.0				1.2	2.2
Estimated division of CO_2 load among buffers										
Hemoglobin, *vol. per cent*				1.0	3.7				0.7	1.4
Phosphate, *vol. per cent*				0.4	1.3				0.3	0.5
Plasma proteins, *vol. per cent*				0.2	0.6				0.1	0.2
Bicarbonate, *vol. per cent*				0.08	0.3				0.06	0.1

blood has lost all of its oxygen (18.5 cc.). In the course of this loss it has absorbed 0.8 as much CO_2, or 14.8 cc.

In table 5 some of the data indicated by measurements on figures 5 and 6 have been put into numerical form. In the "estimated division of CO_2 load among buffers" at the bottom of the table, the estimates are made on the assumption that the relative CO carrying powers of the different buffers of the blood are as represented in table 4 (p. 160). As stated previously, these estimates, especially that concerning the phosphate, are derived from data of only approximate accuracy, but with the possible exception of the phosphate, they may be taken as indicating roughly the relative parts that these factors play in carrying the blood's total load of CO_2.

It is seen from the charts and table 5 that the known factors represented in the curve *AC* permit both bloods to take up normal amounts of CO_2 with normal reaction changes, the blood of J. S. H., taking up 5 volumes per cent of CO_2 with an increase of 0.03 in pH that of J. J. with an increase of 0.01. The CO_2 tension increases are also normal, 6.8 and 5.0 mm. It therefore seems probable that we are acquainted with all the CO_2 carriers of importance.

The point *D* indicates the rise in CO_2 that, had no hemoglobin reduction occurred, would have resulted within the pH change which actually permitted 14.8 volumes per cent CO_2 increase. The vertical difference in the levels of *C* and *D* therefore indicates the volumes per cent of CO_2 carrying capacity that are due to the change from the more acid oxyhemoglobin to the less acid reduced hemoglobin and *vice versa*.

The extent that the buffer factors, other than the change $HbO \rightarrow Hb$, play in enabling the blood to absorb CO_2 within the defined pH limits, is indicated on the charts. It will be noted that, not only is hemoglobin, by virtue of its loss in acidity on reduction, responsible for the greater part of the CO_2 absorption, but also that in carrying the remainder of the CO_2 by buffers, hemoglobin as a buffer plays the major part. The part of hemoglobin in the absorption of the physiologically maximum amounts of CO_2, indicated by the full distance of the curve *AC*, is shown in table 6.

It is thus estimated that in the blood of J. S. H. 84 per cent of the total CO_2 carrying power is due to hemoglobin, in that of J. J. 94 per cent.

It is significant that the blood of J. J., in which hemoglobin carried 94 per cent of the CO_2, absorbed its CO_2 quota over a smaller pH range than did the blood of J. S. H. The two facts are necessarily related. The CO_2 carriers, other than the reaction by which oxyhemoglobin

changes to reduced hemoglobin and thereby frees alkali to form $BHCO_3$, are simply buffers which give up alkali only when the reaction becomes more acid, and in proportion as it does so. If a blood should absorb CO_2 without any change in pH, which is not theoretically impossible, the sole carrier would necessarily be hemoglobin, because it is unique in its ability to provide alkali without being forced to do so by a rise in hydrogen ion concentration.

TABLE 6

	BLOOD OF J. H. S.		BLOOD OF J. J.	
	vol. per cent	*per cent of total*	*vol. per cent*	*per cent of total*
Total CO_2 absorbed	14.8	100	14.8	100
By HbO-Hb change	8.8	59	12.6	85
By hemoglobin as buffer	3.7	25	1.4	9
Total absorption due to hemoglobin	11.5	84	13.0	94

The calculations are, especially for J. J.'s blood, subject to certain approximations. We do not know whether the hemoglobin content of J. J.'s blood was the same as that of J. H. S. To the extent that it was different, our *AC* curve for J. J. is subject to proportional variation. Also we are not certain that 7.35 is the exact pH value normal for J. J.'s blood. Although in a given individual the pH appears to be very constant, in different individuals, according to present data, the pH may vary from 7.25 to 7.45 with 7.35 as the average. If J. J.'s arterial blood were normal at pH 7.25 instead of 7.35, the relative parts played by his different CO_2 carriers would vary less from those of the blood of J. S. H. It is regrettable that no figures are available for the purpose of our calculations which include all the desirable data on the same blood. In the case of the blood of J. S. H., however, it is fairly certain that the range over which we have estimated the changes is normal. At pH 7.35 J. S. H.'s blood shows an estimated CO_2 tension of 41.5 mm., which is about what J. S. H. considers normal for his thoroughly studied blood. Also this is the identical blood on which was based chiefly the calculation of Peters' equation used by us for estimating the effect of reduction of the hemoglobin, so that there appears to be slight danger of error in the estimation of this factor. Of the two sets of data, therefore, it appears that those on the blood of J. S. H. more certainly represent the exact conditions in the organism than do those of J. J., although it is quite possible that the latter are exact also, and the differences between the two bloods represent genuine individual variation.

L60a
2102

CONCLUSIONS

The carriers of carbon dioxide may be most simply described as substances which hold in combination alkali, of which they supply to CO_2, as it enters the blood, sufficient to bind nearly all of it as alkali bicarbonate. When the CO_2 leaves the blood, they recombine with their alkali. By such reactions, despite changes in CO_2 content and CO_2 tension, the alkalinity of the blood in a given normal individual is maintained constant within marvelously narrow limits.

The chief carbon dioxide carrier of the blood is the hemoglobin. It is almost as completely responsible for the transport of carbon dioxide in the blood as it is for the transport of oxygen. Of the alkali furnished to neutralize the CO_2 that enters the venous blood, from 80 to 95 per cent, sometimes possibly all, comes from the hemoglobin. Of such alkali, the greater part is set free when the relatively strong acid, oxyhemoglobin, is changed by loss of oxygen to the much weaker one, reduced hemoglobin. A smaller part of the alkali is furnished by the unchanged oxyhemoglobin, which is itself an efficient buffer at blood reaction. The relatively slight remainder is furnished by the other buffers, viz., the phosphate in the cells, the proteins in the plasma, and the bicarbonate.

The hemoglobin from its location in the cell is able to increase the alkali content of the plasma by withdrawing Cl into the cells from the plasma NaCl, thereby leaving the Na to form $NaHCO_3$. The cells are also freely permeable to the H_2CO_3, so that the latter can enter them to combine with the alkali furnished by the hemoglobin, thereby decreasing the H_2CO_3 of the plasma at the same time that the Cl shift increases its $NaHCO_3$. By means chiefly of the exchange of carbonic and hydrochloric acids with the plasma, the cells are enabled to use their reserves of buffer alkali to maintain constancy of reaction in the plasma, although the buffers themselves (hemoglobin and phosphates) do not leave the cells.

BIBLIOGRAPHY

ABDERHALDEN, E.: Zeitsch. Physiol. Chem., 1898, xxv, 88.

BOHR, C.: Skand. Arch. Physiol., 1905, xvii, 104; Nagel's Handb. Physiol. Menschen., Braunschweig, 1909, i, 55.

BUCKMASTER, G. A.: Journ. Physiol., 1917, li, 164.

CAMPBELL, J. M. H., AND POULTON, E. P.: Journ. Physiol., 1920, liv, 152.

CHRISTIANSEN, C., DOUGLAS, C. G., AND HALDANE, J. S.: Journ. Physiol., 1914, xlviii, 244.